7-7-2020

Characterization of Novel Cellular Systems Using Labelfree Quantitative Proteomics

Robert Gombar
University of Windsor

CHARACTERIZATION OF NOVEL CELLULAR SYSTEMS USING LABEL-FREE QUANTITATIVE PROTEOMICS

By

Robert Gombar

A Dissertation
Submitted to the Faculty of Graduate Studies
through the Department of Chemistry and Biochemistry
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy
at the University of Windsor

Windsor, Ontario, Canada

2020

Characterization of Novel Cellular Systems Using Label-Free Quantitative Proteomics

by

Robert Gombar

APPROVED BY:

_____

J.A Caruso External Examiner
Wayne State University

_____

T.E. Pitcher
Great Lakes Institute for Environmental Research

_____

S. Pandey
Department of Chemistry and Biochemistry

_____

S. Ananvoranich
Department of Chemistry and Biochemistry

_____

P.O. Vacratsis, Advisor
Department of Chemistry and Biochemistry

May 13, 2020

DECLARATION OF CO-AUTHORSHIP/PREVIOUS PUBLICATIONS

I. Co-Authorship Declaration

I hereby declare that this thesis incorporates material that is result of joint research, as follows:

Chapter 2 – Collaboration of Robert Gombar with Dr. Aaron J. Steevensz as co-first authors, Dr. Roland Vergilino, and Dr. Melania E. Cristescu (McGill University) under the supervision of Dr. Panayiotis O. Vacratsis from the University of Windsor. A.J.S developed the experimental procedures and performed all MS analysis while R.G assisted in sample preparation for MS analysis. R.G also processed all MS data and performed bioinformatic analysis. R.G prepared the figures and wrote the manuscript with assistance from A.J.S. The manuscript was edited and revised by P.O.V and M.E.C.

Chapter 3 – Collaboration of Robert Gombar (First author) with Jason A. Lewis, and Dr. Janeen Auld under the co-supervision of Dr. Trevor E. Pitcher (GLIER) and Dr. Panayiotis O. Vacratsis from the University of Windsor. The original hypothesis was designed by T.E.P. All biological samples were captured and acquired by J.A.L and T.E.P. R.G developed all experimental procedures for MS analysis with assistance from P.O.V. MS analysis was primarily performed by J.A. All MS data was processed by R.G. Statistical and bioinformatic analysis was exclusively done by R.G. The complete manuscript and figures was prepared by R.G with edits and revisions done by T.E.P and P.O.V.

Chapter 4 – Collaboration of Robert Gombar with Mehdi Eshraghi as co-first authors, and Yves De Repentigny under the co-supervision of Dr. Rashmi Kothary from the University of Ottawa Centre of Neuromuscular Disease and Dr. Panayiotis Vacratsis from the University of Windsor. The original hypothesis was designed by R.K. All specimens were cared for and handled by Y.D.R and M.E. Synaptosomes were enriched for and prepared by M.E (University of Ottawa) and sent to the University of Windsor for MS analysis. R.G and P.O.V designed and implemented all experimental procedures and performed all MS analysis. All data acquired by MS was analyzed and processed by R.G. Immunohistochemistry and western blot analysis were performed by M.E. Statistical and bioinformatic analysis was done by R.G. The introduction for the manuscript was prepared by M.E and R.K. While the results and discussion section was primarily written by R.G. All

figures regarding MS analysis was prepared by R.G. The manuscript was edited and revised by R.K and P.O.V.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

II. Declaration of previous publication

This thesis includes three original papers that have been previously published for publication in peer-reviewed journals, as follows:

Thesis Chapter	**Publication title/full citation**	**Publication status**
Chapter 2	‡Steevensz A, ‡Gombar R, Vergilino R, Cristescu ME, Vacratsis PO., Proteomic Profile of Daphnia pulex using Data-Independent Acquisition Mass Spectrometry and Ion Mobility Separation. *Proteomics*. (**2018**)18(16). ‡Co-authorship	***Published***
Chapter 3	Gombar R, Pitcher TE, Lewis JA, Auld J, Vacratsis PO., Proteomic characterization of seminal plasma from alternative reproductive tactics of Chinook salmon (Oncorhynchus tswatchysha). *J Proteomics*. 22;157 (**2017**) 1-9.	***Published***
Chapter 4	‡Eshraghi M, ‡Gombar R, De Repentigny Y, Vacratsis PO, Kothary R., Pathologic Alterations in the Proteome of Synaptosomes from a Mouse Model of Spinal Muscular Atrophy. *J Proteome Res*. 2;18(8) (**2019**) 3042-3051. ‡Co-authorship	***Published***

I certify that I have obtained a written permission from the copyright owner(s) to include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Proteomics involves the systematic identification of the protein complement within an array of biological systems. Quantitative proteomics is an extension of proteomics that aims to characterize the changes in protein abundance between different sample states. Mass spectrometry, which measures the mass of ionized molecules in the gas phase, is the predominant analytical tool featured in quantitative proteomic studies. As more genomes are becoming annotated and publicly available, mass spectrometry-based approaches to proteomics have increased in feasibility. Furthermore, the orthogonal detection of all precursor ions and precursor ion fragments, known as data independent acquisition mass spectrometry, has allowed for high-throughput instruments to identify low and high abundant proteins without bias. Data-independent acquisition, in combination with ion mobility, has encouraged the enhancement of protein resolution by further separating ions based on their size, shape and charge. Together, the technological innovations of today's mass spectrometers and advancement of genomic libraries has extended the boundaries for deeper proteome coverage.

The option of using label-free standards for quantitative applications in proteomics is a non-invasive and cost-effective method for measuring protein abundance. However, early label-free strategies suffered from poor resolution and sensitivity issues when analyzing complex mixtures during quantitative studies. By optimizing the use of data-independent acquisition mass spectrometry and ion mobility separation, label-free strategies have joined the toolbox of reliable proteomic platforms for conducting quantitative analysis.

In the following thesis, we present the successful application of using data-independent acquisition mass spectrometry employing orthogonal ion mobility separation to three unique biological systems including *Daphnia pulex* with the supplementation of

using label-free quantitative techniques to explore the protein alterations in the seminal plasma of Chinook salmon, and the synaptosomes of a SMA mouse model. For the *Daphnia pulex* profiling study, our optimized methods have progressed sample preparation methods for the *Daphnia* model system and suggests label free mass spectrometry techniques will be applicable in utilizing the *Daphnia* model in assays that monitor aquatic health dynamics. With regard to the Chinook salmon project, we elucidated statistically significant protein abundance differences between hooknose and jack male tactics. Proteins involved in membrane remodeling, proteolysis, hormonal transport, redox regulation, immunomodulation, and ATP metabolism were among the proteins reproducibly identified at different levels and represent putative factors influencing sperm competition between jack and hooknose males. This study represents the largest seminal plasma proteome from teleost fish and the first reported for Chinook salmon. Lastly, Label-free quantitative proteomics on isolated synaptosomes from spinal cords of a SMA mouse model identified 2030 protein groups. Statistical data analysis revealed 65 specific alterations in the proteome of the central synapses at the early onset stage of disease. Functional analysis of the dysregulated proteins indicated a significant enrichment of proteins associated with mitochondrial dynamics, cholesterol biogenesis, and protein clearance. These pathways represent potential targets for therapy development with the goal of providing stability to the central synapses, thereby preserving neuronal integrity in the context of SMA disease.

I dedicate this thesis to my wife Kayla (Moja lúbavi) and to my mamuška, tajko, sestricka, and mom and dad Lessard.

For this would have all been impossible without their love, support, and devotion.

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my thanks and gratitude to Dr. Panayiotis O. Vacratsis for allowing me the great opportunity and privilege to work in his laboratory at the University of Windsor. Over the years, his guidance and patience has helped foster my abilities to improve as a researcher. More importantly, I have matured to the person I am today thanks to his friendship, and to the life-lessons I learned while working in his lab.

To the members of the Vacratsis lab, both past and present, I would like to thank you for your friendship and support throughout the years. Specially, I would to thank Justin Roberto, Dr. Quidi Geng, Dr. Besa Xhabija, Jas Sohi, Ashley Dadalt, and Christopher Slikboer whose friendships have deeply resonated with my experience in the lab and in life. I personally would like to thank Dr Aaron Steevensz for his assistance during my first year at the Vacratsis lab, which helped provide a foundation for my transition from undergraduate to graduate school. In addition, I would like to thank his passion for science and mass spectrometry – it has been an inspiration for my career in research. To the other students in the department that I have built close friendships to such as: Mitchel DiPasquale, Michael Nguyen, Brett Rickeard, Dave Ure, Anupom Roy, Scott Smith, Scott Roscoe, Cody Caba, Paul Meister, Angela Awada, Amaka Onukwue, and Basel Mansour,I would like to thank you all for enriching my experience at the university.

Most notably, I would also like to extend my thanks and appreciation to my committee members Dr. Sirinart Ananvoranich, Dr. Siyaram Pandey, and Dr. Trevor Pitcher. The advice and suggestions you all provided over the years have helped me progress throughout my academic career and improve as a student. Furthermore, I would like to thank Dr. Trevor Pitcher for allowing me to join his lab and participate in field work -it was an experience I will not forget. Moreover, I would like to thank the amazing secretarial staff in

the department of chemistry and biochemistry Marlene Bezaire, Jayne Pearce, Elizabeth Kickham, and Cathy Wilson. To Marlene Bezaire, thank you for not giving up on me and believing I could finish writing my dissertation. I would like to thank the lab-coordinators in chemistry and biochemistry Norah Franklin, Nedhal Al-Nidawy, Una Lee, and Ronan San Juan for their efforts in maintaining the undergraduate labs and also for also for their support. I would like to thank Joe Lichaa for all the help over the years with technical support and also for the conversations over coffee. To Janeen Auld, thank you for all the help regarding everything to do with mass spectrometry – I would not be able the things I can do today if it weren't for your support. And lastly to Dr. Lana Lee, thank for all the friendly conversations and taking the time listen and encouraging me along my academic path.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

α2M	alpha-2 microglobulin
AI	apolipoprotein I
AII	apolipoprotein II
AIV	apolipoprotein IV
AC	alternating current
AChR	acetylcholine receptor
ADP	adenosine diphosphate
AK1	adenylate kinase 1
AK2	adenylate kinase 2
AK3	adenylate kinase 3
ALS	Amyotrophic lateral sclerosis
ANOVA	analysis of variance
AnxA2	annexin-A2
ApoE	apolipoprotein E
ApoF	apolipoprotein F
ATP	adenosine triphosphate
BOD	biochemical oxygen demand incubator
C3	complement component C3
C3–3	complement component C3-3
C3–4	complement component C3-3
C6	complement component C6
C8α	complement component C8-alpha
C8β	complement component C8-beta
C9	complement component C9

CI	chemical ionization
CID	collision induced dissociation
CMT2D	Charcot-Marie-Tooth type 2D
CNS	central nervous system
COX4	cytochrome c oxidase 4
CRM	charge residue model
DAVID	the database for annotation, visualization and integrated discovery
DC	direct Current
DDA	data dependent acquisition mass spectrometry
ddH2O	doubly distilled water
DIA	data independent acquisition mass spectrometry
DIPP3	diphosphoinositol polyphosphate phosphoryhydrolase 3-alpha
DTT	dithiothreitol
E3 KEG	E3 ubiquitin-ligase KEG-like
e	charge in coulombs
EDTA	ethylenediaminetetraacetic acid
EI	electron Ionization
ELISA	enzyme linked immunosorbent assay
EM	electron microscopy
E_P	electric potential energy
ESI	electrospray ionization
FDR	false discovery rate
FLAMES	field Laboratory for the Assessment of multiple ecological stressors medium
FT-ICR	fourier transform-ion cyclotron resonance
FTLD	frontotemporal lobar degeneration

GFAP	glial fibrillary acidic protein
GlyRS	glycyl tRNA ligase
GM	gradient medium
GO	gene ontology
HCl	hydrochloric acid
$HDMS^E$	ion mobility separation multiplexed acquisition of high and low energy scans
HE	high energy scan
Hi3	label-free internal peptide standard ®
HPLC	high performance liquid chromatography
IDHc	cytoplasmic isocitrate dehydrogenase
IEM	ion evaporation model
IHC	immunohistochemistry
IM	ion mobility
IM-MS	ion mobility mass spectrometry
IP	immunoprecipitation
ITRAQ	isobaric tag for relative and absolute quantitation
K	lysine
K_E	kinetic Energy
KEGG	kyoto encyclopedia of genes and genomes
L	length
LC	liquid chromatography
LDH	lactate dehydrogenase
LE	low energy scan
LFQ	label-free quantitation
m	mass
MALDI	matrix assisted laser desorption ionization

MBP	myelin basic protein
MS1	precursor ion scan
MS2	product ion scan
MS^E	multiplexed acquisition of high and low energy scans
MS/MS	tandem mass spectrometry
m/z	mass-to-charge ratio
mtDNA	mitochondria deoxyribonucleic acid
NADPH	nicotinamide adenine dinucleotide phosphate
NCDN	neurochondrin
NH_4HCO_3	ammonium bicarbonate
NMJs	neuromuscular junctions
PBS	phosphate buffer saline
PDE	phosphodiesterase
PIP3	phosphatidylinositol (3,4,5) triphosphate
PSA	puromycin sensitive aminopeptidase
PTMs	post translation modifications
PSD-95	postsynaptic density protein 95
PVDF	polyvinylidene fluoride
q	total Charge
QqQ	triple quadrupole
R	arginine
RF	radio frequency
SDHA	succinate dehydrogenase A
SDS	sodium dodecyl sulfide
SDS-PAGE	sodium dodecyl sulfide polyacrylamide gel electrophoresis
SHBG	sex hormone binding globulin

SILAC	stable isotope labeling of amino acid in cell culture
SMA	spinal muscular atrophy
SMA-DV	distal SMA type-V
SMN1	survival motor neuron 1
snRNPs	small nuclear ribonucleoproteins
SOD	superoxide dismutase
SRIG	stacked ring ion guide
SV2	synaptic vesicle glycoprotein 2
t	time
TBST	tris buffered saline with tween 20
TK2	thymidine kinase 2
TMEM106B	transmembrane protein 106B
TMT	tandem mass tag
TOF	time-of-flight
Tris-HCl	tris(hydroxymethyl)aminomethane hydrochloride
TWIMS	travelling wave ion mobility separation
UPLC	ultra performance liquid chromatography ®
UPS	ubiquitin-proteasome system
v	velocity
VGlut1	vesicular glutamate transporter 1
V_S	accelerating voltage
WB	western blot
z	elementary Charge

CHAPTER 1- LITERATURE REVIEW

1.1 Proteomics – Interrogating the Proteomes of Living Systems

The fundamental task in proteomics is to acquire a comprehensive understanding of the protein levels and their functional correlation to the physiology of a particular biological system [1]. In brief, the proteome is defined as the total protein complement, which consists of the entire set of proteins that can or will be expressed within a system [2]; the system being either a specific cell line, tissue, or whole organism [3].

In the context of biological systems, the proteome is remarkably complex in regard to its composition and function [4]. As such, there are two categories of proteomic research that has warranted the attention of the life sciences: expression proteomics, and functional proteomics [5]. The former involves the qualitative detection of expressed proteins in addition to the quantitative analysis of their abundances in order to reveal important biomarkers. Naturally, expression proteomics aims to characterize the changes in protein expressions between two or more states; usually the up-regulation or down-regulation of proteins between a physiologically healthy phenotype and a phenotype that resembles the pathology of disease progression [6]. In contrast, the changes in protein expression can also be useful in characterizing the importance of evolutionary advantages from species that employ different life history traits (e.g. alternate reproductive tactics) [7]. The latter approach aims to characterize protein activities in a spatial-temporal manner by investigating protein-protein interactions that form multi-protein complexes, changes in protein activity through post-translation modifications (PTMs), protein function with respect to subcellular localization, and protein networks involved in signal transduction [8].

To achieve the primary objectives of proteomics, it is important to discuss the existing technologies that enable global analysis of the protein complement, their interactions, abundance, and activity. Some of the conventional sample preparation techniques used in proteomic research comprises the use of Sodium dodecyl sulphate Polyacrylamide gel electrophoresis (SDS-PAGE) to visually compare and detect changes in proteomes. Other techniques invoke the specificity and sensitivity of antibodies such as: Western Blot (WB), immunohistochemistry (IHC), immunoprecipitation (IP), and enzyme linked immunosorbent assay (ELISA) [9]. As powerful as these techniques can be, they are still limited to the analysis of single proteins or directed towards the subcellular proteome. One particular technique, that has brought proteomics to the forefront of systems biology, is high resolution mass spectrometry (MS). Because of its high mass accuracy, sensitivity, and high-throughput capabilities, MS applications have a significant capacity to investigate the dynamic nature of biological systems at the proteome level [10]. Therefore, biological-MS is the ideal platform for hypothesis-driven proteomic research.

1.2 Mass Spectrometry – Instrumentation & Theory

Mass spectrometry (MS) is a distinct branch of science that studies matter at the molecular level – specifically, by measuring the behavior of gas phase ions in the presence of an electric, and/or magnetic field [11]. The mass-to-charge-ratio (m/z) of ions provides valuable insight to their molecular mass, thus enabling MS to identify the chemical signature and composition of molecular structures [12].

In regard to instrumentation, all mass spectrometers will consist of three core elements: The ionization source; the mass analyzer; and the detector. The purpose of the ionization source is to generate ions from any analyte present in a sample, and also to assist in the transmission of ions to the mass analyzer through their desorption into the gas phase [13]. The function of the mass analyzer is to help guide and sort the ions, based on their m/z value, by focusing their trajectories using either, or in combination of, an electric field, magnetics field, and field-free sector. In addition to focusing ions, the mass analyzer's secondary function is to perform tandem mass spectrometry (MS/MS) [14]. Lastly, the detector, of mass spectrometer, is integral in the recording of both the presence, and intensity of ions after they have been sorted based on their m/z values [15].

1.2.1 Ionization Source

A fundamental step, that is critical to the success of MS analysis, is the ability produce charged particles in the gaseous state. Although all mass spectrometers will have a compartment dedicated to the ionization of analytes, different ionization mechanism can be fundamentally distinct. Some examples of ionization sources include electron ionization (EI), and chemical ionization (CI) [11]. However, the conditions for ionization that EI and CI provide are not conducive for the analysis of molecules that are derived from biological

systems such as proteins, DNA/RNA, carbohydrates, lipids, and metabolites [16]. Two prominent ionization sources that are useful for the analysis of biomolecules, especially for proteomic investigations, are Matrix-assisted laser desorption ionization (MALDI), and electrospray ionization (ESI) [17,18].

MALDI is a well-known and powerful method in the production of intact gas-phase ions of biological related molecules. First described by Karas and Hillenkamp in the 1987, MALDI's success has been dependent on the use of a small organic molecule known also as the matrix [19]. The purpose of the matrix is three-fold. The first is to provide a chemical environment that is compatible with that of the analytes all while being chemically inert. In addition, the matrix affords favourable conditions for the formation of charged products -either negatively or positively charged. The second involves the physico-chemical properties of the matrix itself which is the ability to become energetically excited by photoionization. And lastly, the matrix provides a protective layer which buffers the analyte from the destructive properties of the incoming UV-laser [20]. The process of ionization itself is dependent on the formation of a solid deposit that is embedded with analytes co-crystalized within matrix molecules. Subsequently, laser irradiation of the solid deposit causes localized rapid-heating of the matrix which initiates the process of sublimation of the matrix alongside the analyte molecules into the gas-phase [21]. Although the process of ion formation is not fully understood, there are two main mechanisms that have been described for MALDI. Gas-phase protonation is the first mechanism proposed which states that analytes become ionized through charge transfer as they become desorbed into the gas-phase [22-24]. The competing theory is known as the lucky-survivor model. This model states that the analytes are initially charged while in the

condensed state all while being surrounded by their respective counter-ions. During the sublimation process, the matrix and counter ions are stripped from the analytes [25,26]. However, both the lucky-survivor model and gas-phase protonation model have been implicated in the formation of singly charged ions.

The process of ionization for ESI is fundamentally different from MALDI. While MALDI generates singly charged ions from the solid-state, ESI produces multiply charged ions starting from a liquid medium. The composition of the solution is comprised of highly conductive solvent with polar character [27]. First described by John Fenn et al., the process of ion formation for ESI begins by subjecting the solution to a strong electric field while passing the sample through a needle/capillary at a low flow rate. By applying an electric potential difference of 3-6KV between the capillary tip and counter electrode, the electric-field induces a charge accumulation at the tip. With the addition of a co-axial sheath gas, the charge accumulation at the tip disperses the solution into highly charged droplets (nebulization). Furthermore, the application of heated gas which flows in opposition to the direction of the spray helps evaporate and any residual solvent resulting in the release of newly formed ions [28,29]. Essentially, the over-accumulation of like-charges creates repulsive forces inside the droplet in order to exceed the surface tension of the solvent leading to coulombic explosion of the droplet itself and desorption of the ions into the gas-phase. This process of ion formation is known as the ion-evaporation method (IEM) and favors the production of ions of low m/z values [30]. The charge residue model (CRM), on the other hand, favors the formation of ions with high m/z. The competing model states that the formation of ions occurs through a process of solvent evaporation which ultimately results in an ion's desorption from a singular droplet [31].

1.2.2 Mass Analyzers:

More of a filter rather than participating in the analysis of ions, mass analyzers play an important role in the separation of ions based on their m/z values . With the exception of a magnetic sector, most mass analyzers operate under the principle that ions can be manipulated in space by utilizing electric fields [14]. Some common mass analyzers that are used in most mass spectrometers include quadrupoles, quadrupole ion-trap, time-of-flight (TOF), and Fourier transform ion cyclotron resonance (FT-ICR) [11,14]. However, for simplicity and the purpose the of the present document, only quadrupoles and TOF will be reviewed in the following section.

A quadrupole mass analyzer, as the name implies, consist of four identical rods that are oriented in space to produce a symmetrical region of oscillating electric field [32]. The oscillating electric field is generated by applying a voltage from a constant direct current (DC) with the addition of a superimposed radio frequency (RF) voltage of alternating current (AC) to each respective pair of rods. However, the electric potential difference that is supplied by the DC will be positive for one pair of rods and negative for the other, x-plane and y-plane, respectively **[Fiure 1.1]** [33]. In addition, the RF for the AC voltages will be 180° out of phase between the opposing pair of rods. The significance of combining both DC and AC ensures that, during the initial scanning mode, the flight path of only select few ions of certain m/z values will survive the passage along the quadrupole and become detected. Meanwhile, the rest of the ions, whose flight paths do not resonate with the selected DC and AC voltages, will discharge on the poles [2,11,14,33]. Transmission mode, on the other hand, operates only on RF voltages to ensures a wider mass window of ions with a varying range of m/z values -which has been extremely valuable for the

identification proteins regarding proteomic investigations [2,11]. While quadrupole mass analyzers have robust focusing power, they are only able to resolve ions within a narrow mass range (4000Da mass limit) [11,14]. This limitation is primarily circumvented by exploiting the natural characteristics of ESI and the production of multiply charged ion species [11].

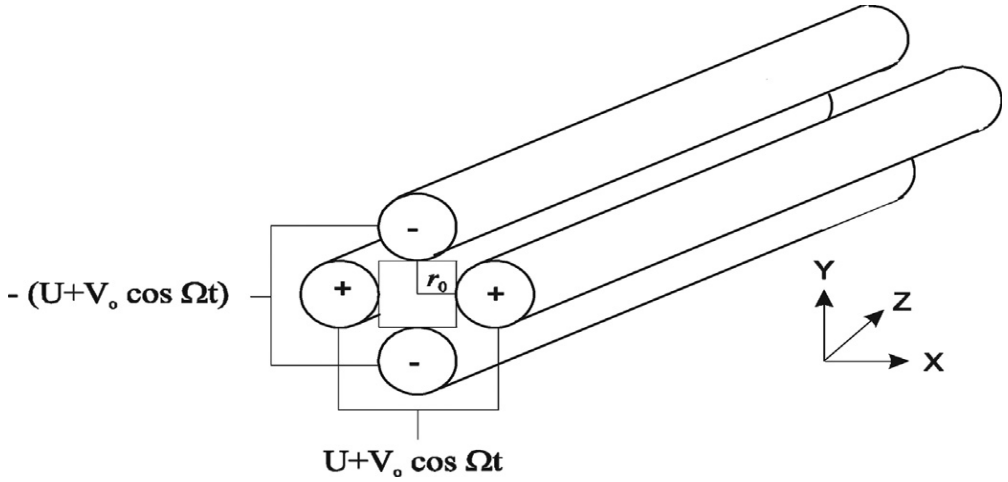


Figure 1.1 – Diagram of a quadrupole mass filter. As ions enter the quadrupole mass filter, they will traverse along the z-plane while being subjected to an oscillating electric field imposed by the rods in the x and y positions. U and V_o represent the DC current and RF voltage (AC), respectively. From March et al., *Int. J. Mass. Spectrom*, 377 (2015) 316-328 **[33]**.

The principles of ion separation regarding TOF mass analyzers involve the use of a homogenous electric field which propagates ions into a field-free region of the mass spectrometer known as the drift-tube; the drift-tube is kept under high vacuum/ low pressure conditions in order to avoid unwanted collision of ions [34]. In the drift-tube, there will be a distribution of ions of various masses that are characterized on the differences in velocities. Due to the homogenous electric field, each ion will acquire the same kinetic energy (K_E), however, the velocities of ions will differ on the basis of their m/z values [11,34]. For instance, ions with a smaller m/z value will travel faster in the drift-tube and will be detected earlier than ions with a larger m/z value. Since K_E is related to electric potential (E_p) [Equation 1], the time of when the ion will be detected can be calculated based on their m/z values. Where *m* is mass, *v* the velocity, *q* is the total charge of the ion, and V_s is the accelerating voltage applied to the ions [11].

$$\frac{1}{2}mv^2 = qV_s \quad \text{.....Equation 1}$$

To solve for velocity, the total charge can be expanded into its elementary charge *z*, and the charge of the ion in coulombs *e* [Equation 2] to yield [Equation 3].

$$q = ez \quad \text{.....Equation 2}$$

$$v = \frac{2ezVs^{1/2}}{m} \quad \text{..... Equation 3}$$

Since ions are accelerated into a drift-tube of known length (L), the time (t)that is taken to cover the distance is determined by substituting velocity (*v*) into [Equation 4].

$$t = \frac{L}{v} \text{.....Equation 4}$$

Because the length, accelerating voltage, and charge is held constant for the instrument, the time that is required for the ion to be detected is relate to the ion's mass and charge [Equation 5].

$$t^2 = \frac{m}{z} \left(\frac{L}{2eVs}\right) \text{.....Equation 5}$$

1.2.3 Ion Mobility Mass Spectrometry

First introduced in the 1960s, ion mobility MS (IM-MS) is a technique that separates ions in the gas phase based on their size, shape, and charge [35]. Since the objectives of proteomics is to characterize the total proteome, IM-MS has been proven to be a powerful method in improving peptide resolution and thus increasing the dynamic range of proteome coverage in a high throughput manner [36]. Travelling wave IM separation (TWIMS) is a relatively novel methods in IM separation which is incorporated within the SYNAPT series of mass spectrometers that are manufactured by Waters [37].

The conventional set-up and operation of TWIMS is dependent on the stacked ring ion guides (SRIG), as depicted in **[Figure 1.2]**. As illustrated, RF voltages are oriented in alternating fixed intervals with respect to the SRIG with the addition of a superimposed DC voltage to provide directionality of the ions [38]. The principle of separation involves the production of symmetrical wave-like motion of ions based on the RF electric field. In addition, the IM cell will be held under atmospheric conditions with inert gas flowing counter to the movement of ions. As ions surf the wave, they will experience a drag force due to collisions with the gas. The greater the collision cross-section of the ion, the longer it takes for ions to travel the IM drift-tube. For instance, ions with a larger size and surface area will experience greater amount of drag force than more compact molecular ion with the same m/z [35-38]. As a result, peptide and protein ions that co-elute during the chromatography step can be separated by IM.

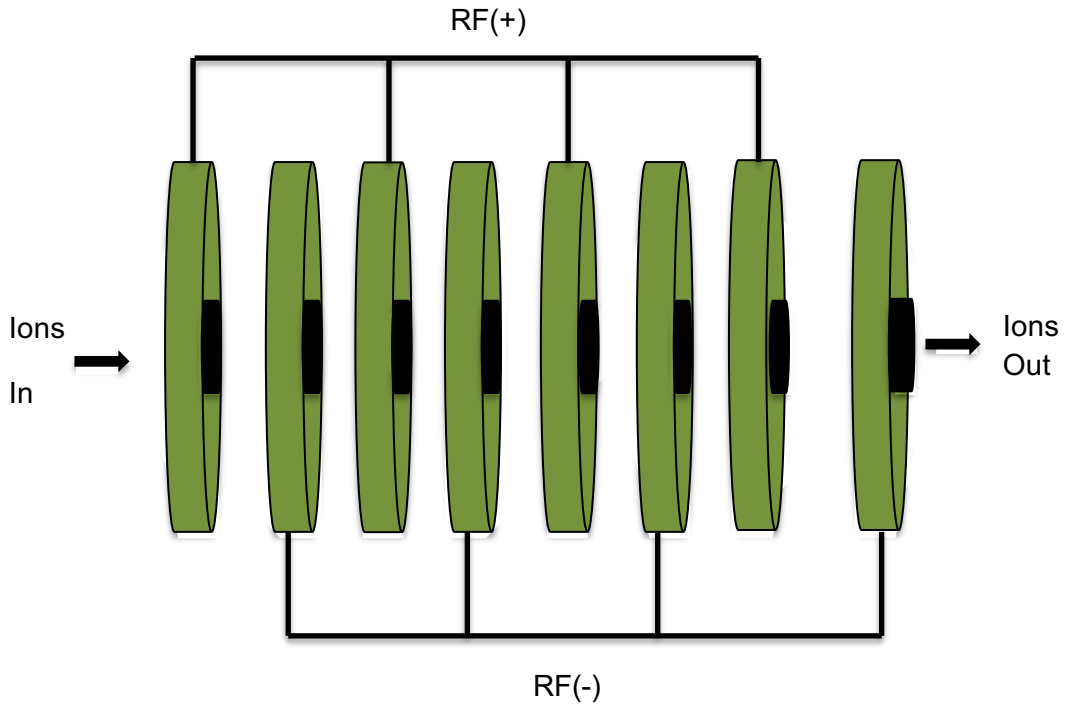


Figure 1.2 – Stacked ring ion guide of travelling wave ion mobility cell. Adapted from Cumeras et al., Analyst, 140:5 (2015) 1376-1390 **[38]**.

1.3 MS-Based Proteomic Workflows

The proteomic pipeline begins with isolation of biological material followed by sample preparation for MS, MS analysis, data-processing, and lastly ending with bioinformatic analysis **[Figure 1.3]** [39]. Often overlooked, a very important step in proteomic workflows is the reproducible isolation of biological material [40]. The proteome itself is dependent on the total number of protein-coding genes [41]. However, the complexity of the proteome can be modulated by several factors that act on separate junctures during the process of protein expression. For instance, the abundance of proteins can be affected by copy number variation of genes [42]. In addition, genes can be regulated either through their amplification or silencing which can result in tissue-specific protein expression [43]. At the transcriptome level, post-transcriptional machinery can modify mRNA by-products through a process of alternative splicing resulting in the production of multiple proteoforms from a single gene [44-46]. Lastly, at the proteome level, a protein's function, localization, and lifespan can be dictated upon the addition or removal PTMs [47]. Furthermore, the stoichiometries of the protein landscape is affected in spatial-temporal manner, meaning that not all proteomes will be alike under strict conditions [48]. This is especially true in regard to the equilibrium between protein synthesis and protein degradation [49]. Thus, it is imperative that appropriate biological sources are chosen when evaluating proteomic hypotheses using MS.

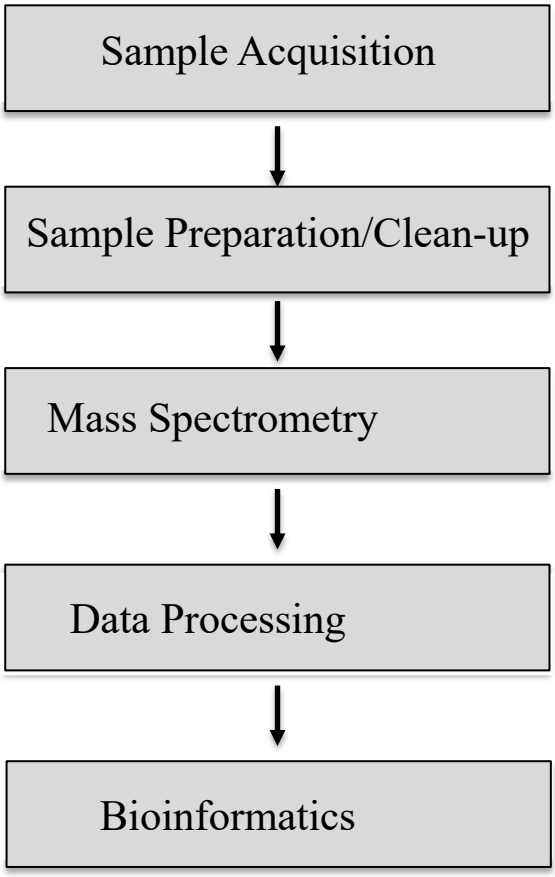


Figure 1.3 – Typical flow-diagram for proteomic workflows. Adapted from Malmström et al., *Curr Opin Biotechnol*, 18:4 (2007) 378-384 **[39].**

1.3.1 Bottom-Up Proteomics

The so-called bottom-up/shotgun approach to proteomics involves the analysis of proteins by measuring peptides generated from the proteins that have been subjected to proteolytic digestion. Although this may seem counter intuitive, enzymatic digestion of proteins are necessary to normalize a heterogenous complex mixture of proteins that allows for easier identification based on tandem mass spectrometry [50]. In comparison, top-down proteomic studies involve the MS measurement of fully intact proteins and has been successful in the identification of hundreds of different proteins as well as mapping various PTMs. However, top-down proteomics is usually applied in specialized cases (e.g. histone proteomics) while bottom-up proteomics is the more common approach for typical quantitative proteomic applications [52-55].

A typical procedure for bottom-up proteomics is dependent on an a-priori knowledge of the total protein database for which protein identification will be inferenced [56]. For instance, after the proteins have been digested, separated using on-line liquid chromatography, and analyzed using MS, peptide identification is achieved by matching the protein mass fingerprint of the empirical mass spectral data alongside the in-silico digested protein database [50,56]. A common proteolytic enzyme used in most proteomic investigations is trypsin; a serine protease that specifically cleaves proteins at the carboxy terminus of lysine (K), and arginine (R) residues. Some characteristics of trypsin that make it ideal for bottom-up proteomics include its high specificity for K and R residues, its production of peptides of optimal length for MS measurements, its compatibility with downstream sample cleanup steps that are easily streamlined with LC-MS, and the

availability of at least one basic residue (K or R) that can be readily protonated in an acidic environment [57,58].

1.3.2 Tandem Mass Spectrometry

The application of tandem mass spectrometry (MS/MS) is an integral tool for proteomic research [59]. It provides the unambiguous -albeit probability based - method of identification of peptides established on the analysis of the fragment data that is derived from the ordered decomposition of precursor ions (i.e. the ionized form of the full-length peptide) [60].

The classical scheme of MS/MS operates on the basis that there are two separate mass analyzers placed in tandem [61]. In a specialized compartment of the mass spectrometer, termed the collision cell, fragmentation occurs through a process of collision induced dissociation (CID). In comparison to other sections of the mass spectrometer, this specialized chamber is under atmospheric/ low vacuum pressure conditions [62]. This is best exemplified by the use of a triple quadrupole (QqQ). The capitol Q refers to a quadrupole that is operated using both AC and DC voltages to scan for specific m/z values of precursor ions and fragment ions. In contrast, quadrupoles that are controlled strictly using RF mode are designated by the lowercase q - also formally known as the collision cell [63]. Although the Synapt G2-S*i* used in our studies is not a QqQ instrument, it does have a distinct section that is recognized as the Triwave which houses the trap, IM, and transfer cells **[Figure 1.4]** [64]. The trap and transfer cells are comprised of a similar SRIG design to that of the IM cell. The purpose of the trap and transfer cells are to guide/focus ions into the IM cell and TOF regions, respectively. In addition, they can both act as individual collision cells [37,64]. However, the sequence of CID is dependent on the type

instrument mode chosen for proteomic analysis and the relevance of this will be discussed in the following section.

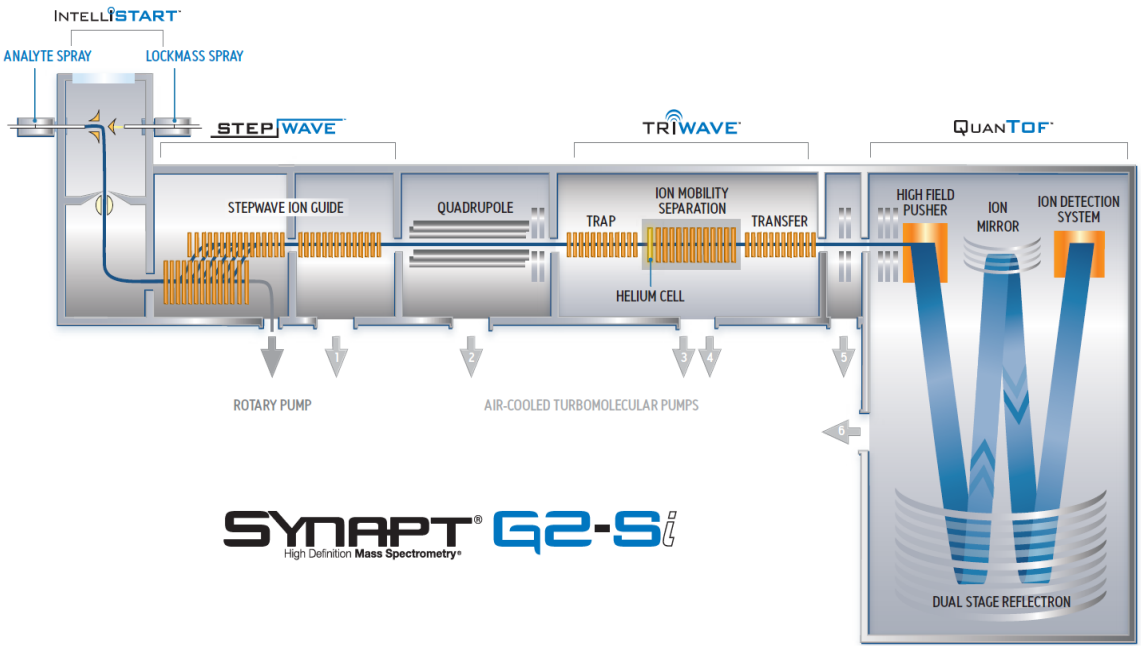


Figure 1.4 – Schematic diagram of the Synapt G2-si mass spectrometry by Waters ®. As illustrated from Ponthus et al., *Int. J. Ion Mobil. Spec*, 16 (2013) 95-103 **[64]**.

With respect to the fragmentation process, an important feature of the collision cell is that is it filled with an inert gas (i.e collision gas). As precursor ions enter this chamber, their momentum is escalated by manipulating the RF voltage strength in order to increase the chances of collisions with the collision gas. During the event of a collision between the precursor ion and collision gas, some of the kinetic energy will become converted into vibrational energy. Once the threshold for vibrational energy is surpassed, the precursor ion will decompose into product ions (i.e fragment ions). For peptides, the precursor ions will fragment at particular nodes within the peptide backbone **[Fig. 1.5]**, and depending on the localization of charge, fragmentation of the precursor ion will yield predominantly y (C-terminal) and b (N-terminal) ions [67].

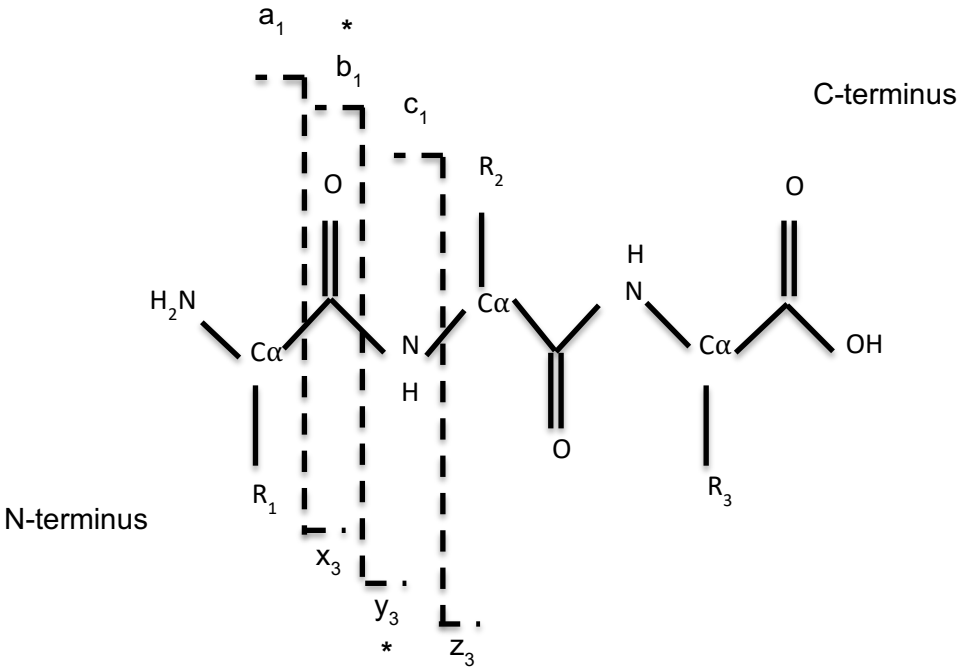


Figure 1.5 - Nomenclature for common ion types produced by CID. The asterisk (*) denotes the most prominent ion types observed. As adapted from Wysocki et al., *Methods*, 35:3 (2005) 211-222 **[67]**.

1.3.3 Data-Independent Acquisition Mass Spectrometry

There are two general data-acquisition strategies that are extremely useful for the comprehensive identification of proteins related to bottom-up proteomic applications. The first, which is a dominant method utilized by most high-resolution mass spectrometers, is referred to as data-dependent acquisition (DDA). In this approach, fragment data (MS2) is dependent on the initial survey scan of precursor ions (MS1) [68]. During MS1, precursor ions are selected in real time for fragmentation based on abundance rank, signal intensity, and charge-state. However, by evading the identification of lower abundant proteins, the main disadvantage of DDA includes the risk that all peptides selected for fragmentation will be derived only from highly abundant proteins [69]. In addition, variation in chromatographic separation of peptides can obscure the sequential fragmentation of precursor ions as the retention times of peptides may differ over time [70]. However, modern DDA mass spectrometers are able to mitigate these performance issues with improvements made to faster scanning times and dynamic exclusion of peptides [69].

The second strategy, known as data-independent acquisition (DIA), helps to mitigate the bias towards the selection of highly abundant peptides and variation in chromatographic separation. The typical procedure for DIA involves the simultaneous analysis of precursor and fragment ions without the prerequisite analysis of abundance and intensity of precursor ions from an MS1 scan [71]. This type of acquisition is more formally known as MS^E, where the term *E* stands for *everything*. While this type of acquisition strategy in comparison to DDA produces chimeric spectra that is very complicated to decipher, there are programs such as PLGS and Progenesis (developed by WATERS) that is able to deconvolute the MS/MS spectra produced by DIA strategies into actual

intelligible information [72]. With respect to the SYNAPT G2-Si, DIA is performed by cyclical CID modes which iterates low energy (LE) and high energy (HE) scans – the LE energy scans are insufficient to fragment precursor ions while the HE scans concomitantly produces fragment ions from the full- length precursor ions; toggling between LE and HE for DIA strategies occurs in the trap cell of the TriWave of the SYNAPT before entering the TOF region [73.74]. To improve resolution and protein identification, IM can be employed in an orthogonal manner prior to DIA to separate precursor ions based on size, charge, and shape. This form of DIA is called high definition MS^E ($HDMS^E$) [75,76].

1.4 Quantitative Proteomics

The protein landscape is constantly fluctuating in response to its internal and external environment. The result of which will require accurate means of quantitating changes to the proteome. Quantitative proteomics is a branch of proteomics that aims to characterize changes in protein abundance for a variety of biological applications such as discovering novel biomarkers associated with diseases or advancing our understanding of fundamental biological mechanisms by measuring the protein alterations in response to extracellular stimuli [77-79].

While MS is a robust and sensitive tool for protein identification, it is not a technique that is inherently quantitative [80]. However, emerging technologies are being used to supplement quantitative measures for MS. Some examples of these methods predominantly use labeling strategies to discern differences between protein abundance such as: stable isotope labeling of amino acid in cell culture (SILAC) [81,82], isobaric tags for relative and absolute quantitation (iTRAQ) [83], and tandem mass tagging (TMT) [84]. SILAC involves the endogenous incorporation of amino acids, of various isotopic masses,

in metabolically active cells [81,82]. Conversely, iTRAQ and TMT are post-lysis labeling strategies that are applied to the peptide level with an affinity for primary amines [83,84]. These techniques can provide quantitative results with high mass accuracy and have the advantage of being able to multiplex several biological replicates into a single run, reducing instrument usage and measurement time [81-84]. There are some disadvantages to labeling strategies as well. For instance, these techniques are dependent on the reproducible labeling of their targets to completion, thus the stochasticity of labeling can add variability to quantitation. In addition, the capability of multiplexing can over-complicate the already convoluted spectra of DIA MS/MS [69].

Currently, label-free quantitation (LFQ) is the only method compatible for DIA-MS. In addition, LFQ is an non-invasive and relatively inexpensive method of quantitation in regard to measuring protein abundance with MS [69]. The earliest form of LFQ includes monitoring ion-count and ion-intensity. The former, also known as spectral counting, predicates on the strong correlation between abundant proteins and the amount of MS/MS spectra they produce. The latter, on the other hand, establishes relative quantitative results of protein abundance based on the ion-intensity which is calculated by the area under the curve, or peak height [69,85,86]. A more reliable form of LFQ, which provides close to absolute quantitative results, is a method termed either Top3 or Hi3. First reported by Silva et al., this type of LFQ utilizes a set of internal peptide standards of known concentration which measures protein abundance by extrapolating the top three ranking abundant peptides in relation to the standard [87].

1.5 Objectives

The development of quantitative proteomic methodology within our laboratory has facilitated collaborative opportunities to advance the knowledge of a variety of fundamental cellular processes and model systems. The objectives of my dissertation are to adapt and implement the utilization of various quantitative proteomic techniques using ion mobility in conjunction with data independent acquisition mass spectrometry to interrogate various cellular systems using label-free quantitative strategies.

Specifically:

1) **Proteomic Characterization of *Daphnia pulex* for Eco-Toxicological Studies.**

2) **Quantitation of Seminal Plasma Protein Abundance Between Different Reproductive Tactics of Chinook Salmon.**

3) **Comparative Proteomic Analysis of Synaptosomes Isolated from Spinal Muscular Atrophy Mice.**

1.6 – REFERENCES

[1] R. Aebersold, M. Mann, Mass spectrometry-based proteomics, Nature 13 (2003) 198-207.

[2] J. Lovric, Introducing Proteomics: from concepts to sample preparation, mass spectrometry and data analysis, 1 ed. Oxford / UK: John Wiley & Sons Ltd (2011).

[3] M.R. Wilkins, J.C. Sanchez, A.A. Gooley, R.D. Appel, I. Humphery-Smith, D.F. Hochstrasser, K.L. Williams, Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it, Biotechnol Genet Eng Rev 13 (1996) 19-50.

[4] J.W. Harper, E.J. Bennett, Proteome complexity and the forces that drive proteome Imbalance, Nature 537 (2016) 328–338.

[5] P. R. Graves, T.A. Haystead, Molecular biologist's guide to proteomics, Microbiol Mol Biol Rev 66 (2002) 39–63.

[6] W.P. Blackstock, M.P. Weir, Proteomics: quantitative and physical mapping of cellular proteins, Trends Biotechnol 17 (1999) 121-127.

[7] R. Gombar, T.E. Pitcher, J.A. Lewis, J. Auld, P.O. Vacratsis, Proteomic characterization of seminal plasma from alternative reproductive tactics of Chinook salmon (Oncorhynchus tswatchysha), J Proteomics 157 (2017) 1-9.

[8] M. Monti, S. Orrù, D. Pagnozzi, P. Pucci, Functional proteomics, Clin Chim Acta 357 (2005) 140-150.

[9] B. Aslam, M. Basit, M.A. Nisar, M. Khurshid, M.H. Rasool, Proteomics: Technologie sand Their Applications, J Chromatogr Sci 55 (2017) 182-196.

[10] X. Han, A. Aslanian, J.R. Yates 3rd, Mass spectrometry for proteomics, Curr Opin Chem Biol 12 (2008) 483–490.

[11] E. Hoffmann, V. Stroobant, Mass Spectrometry: Principles and Applications, 3rd ed. John Wiley & Sons: Chicester, England, (2007).

[12] A. El-Aneed, A. Cohen, J. Banoub, Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers, Appl Spectrosc Rev 44 (2009) 210-230.

[13] H. Awad, M. M. Khamis, A. El-Aneed, Mass Spectrometry, Review of the Basics: Ionization, Appl Spectrosc Rev 50 (2015) 158-175,

[14] A. M. Haag, Mass Analyzers and Mass Spectrometers, Adv Exp Med Biol 919 (2016) 157-169.

[15] S. Medhe, Mass Spectrometry: Detectors Review, Chem Biomol Eng 3 (2018) 51-58.

[16] H.D. Beckey, Field desorption mass spectrometry: A technique for the study of thermally unstable substances of low volatility, Int J Mass Spectrom Ion Phys 2 (1969) 500–503.

[17] P. Millares, E.J. Lacourse, S. Perally, D.A. Ward, M.C. Prescott, J.E. Hodgkinson, P.M. Brophy, H.H. Rees, Proteomic profiling and protein identification by MALDI-TOF mass spectrometry in unsequenced parasitic nematodes, PLoS One 7 (2012).

[18] M. Wilm, Principles of electrospray ionization, Mol Cell Proteomics 10 (2011).

[19] M. Karas, F. Hillenkamp, Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons, Anal Chem 60 (1988) 2299-2301.

[20] I.C. Lu, C. Lee, Y.T. Lee, C.K. Ni, Ionization Mechanism of Matrix-Assisted Laser Desorption/Ionization, Annu Rev Anal Chem (Palo Alto Calif) 8 (2015) 21-39.

[21] M. Karas, D. Bachmann, U. Bahr, F. Hillenkamp, Matrix-assistedultravioletlaser desorption of non-volatile compounds, Int J Mass Spectrom Ion Process 78 (1987) 53–68.

[22] R. Knochenmuss, R. Zenobi, MALDI ionization: the role of in-plume processes, Chem Rev 103 (2003) 441–452.

[23] R. Knochenmuss, (2006) Ion formation mechanisms in UV-MALDI, Analyst 131 (2006) 966–986.

[24] R. Zenobi, R. Knochenmuss, (1998) Ion formation in MALDI mass spectrometry, Mass Spectrom Rev 17 (1998) 337–366.

[25] M. Karas, M. Glu¨ckmann, J. Scha¨fer, Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors, J Mass Spectrom 35 (2000) 1–12.

[26] T.W. Jaskolla, M. Karas, Compelling evidence for lucky survivor and gas phase protonation: the unified MALDI analyte protonation mechanism, J Am Soc Mass Spectrom 22 (2001) 976–988.

[27] S. Banerjee, S. Mazumdar, Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte, Int J Anal Chem (2012).

[28] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules, Science 246 (1989) 64–71.

[29] M. Mann, C.K. Meng, J.B. Fenn, Interpreting mass spectra of multiply charged Ions, Anal Chem 61 (1989) 1702–1708.

[30] J.V. Iribarne, B.A. Thomson, One the evaporation of small ions from charged droplets, J Chem Phys 64 (1976) 2287–2294.

[31] P. Kebarle, A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry, J Mass Spectrom 35 (2000) 804–817.

[32] J.P. Savaryn, T.K. Toby, N.L. Kelleher, A researcher's guide to mass spectrometry-based proteomics, Proteomics 16 (2016) 2435–2443.

[33] R.E. March, J.F.J. Todd, Radio frequency quadrupole technology: Evolution and contributions to mass spectrometry, Int J Mass Spectrom 377 (2015) 316-328,

[34] U. Boesl, Time-of-flight mass spectrometry: Introduction to the basics, Mass Spectrom Rev 36 (2017) 86-109.

[35] G. Ben-Nissan, M. Sharon, The application of ion-mobility mass spectrometry for structure/function investigation of protein complexes, Curr Opin Chem Biol 42 (2018) 25–33.

[36] S. Pfammatter, E. Bonneil, F.P. McManus, S. Prasad, D.J. Bailey, M. Belford, J.J.

Dunyach, P. Thibault, A Novel Differential Ion Mobility Device Expands the Depth ofProteome Coverage and the Sensitivity of Multiplex Proteomic Measurements, Mol Cell Proteomics 10 (2018) 2051-2067.

[37] A.A. Shvartsburg, R.D. Smith RD, Fundamentals of traveling wave ion mobility Spectrometry, Anal Chem 23 (2008) 9689-9699.

[38] R. Cumeras, E. Figueras, C.E. Davis, J.I. Baumbach, I. Gràcia, Review on ion mobility spectrometry. Part 1: current instrumentation, Analyst 140 (2015) 1376–1390.

[39] J. Malmström, H. Lee, R. Aebersold, Advances in proteomic workflows for systems biology, Curr Opin Biotechnol 18 (2007) 378-384.

[40] D.J. Orton, A.A. Doucette, Proteomic Workflows for Biomarker Identification Using Mass Spectrometry - Technical and Statistical Considerations during Initial Discovery, Proteomes 1 (2013) 109–127.

[41] E.A. Ponomarenko, E.V. Poverennaya, E.V. Ilgisonis, M.A. Pyatnitskiy, A.T. Kopylov, V.G. Zgoda, A.V. Lisitsa, A.I. Archakov, The Size of the Human Proteome: The Width and Depth, Int J Anal Chem (2016).

[42] R. Beroukhim, C.H. Mermel, D. Porter, G. Wei, S. Raychaudhuri, J. Donovan, J. Barretina, J.S. Boehm, J. Dobson, M. Urashima, K.T. Mc Henry, R.M. Pinchback, A.H. Ligon, Y.J. Cho, The landscape of somatic copy-number alteration across human cancers, Nature 463 (2010) 899-905.

[43] L. Haery , H. Greulich, M. Reich, W. Winckler, M.S. Lawrence et al, The landscape of somatic copy-number alteration across human cancers, Nature 18 (2010) 899-905.

[44] M.J. Roth, A.J. Forbes, M.T. Boyne 2nd, Y.B. Kim, D.E. Robinson, N.L. Kelleher, Precise and parallel characterization of coding polymorphisms, alternative splicing, and modifications in human proteins by mass spectrometry, Mol Cell Proteomics 4 (2005)1002-1008.

[45] Q. Pan, O. Shai, L.J. Lee, B.J. Frey, B.J. Blencowe, Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing, Nat Genet 40 (2008) 1413-1415.

[46] L.M. Smith, N.L. Kelleher, Consortium for Top Down Proteomics. Proteoform: a single term describing protein complexity, Nat Methods 10 (2013) 186-187.

[47] M. Mann, O.N. Jensen, Proteomic analysis of post-translational modifications, Nat Biotechnol 21 (2003) 255-261.

[48] M. Buszczak, R.A. Signer, S.J. Morrison, Cellular differences in protein synthesis regulate tissue homeostasis, Cell 159 (2014) 242-251.

[49] T. Yu, Y. Tao, M. Yang M, Profiling human protein degradome delineates cellular responses to proteasomal inhibition and reveals a feedback mechanism in regulating proteasome homeostasis, Cell Res 4 (2014) 1214–1230.

[50] Y. Zhang, B.R. Fonslow, B. Shan, M.C. Baek, J.R. Yates 3rd, Protein analysis by shotgun/bottom-up proteomics, Chem Rev 113 (2013) 2343–2394.

[51] A.D. Catherman, O.S. Skinner, N.L. Kelleher, Top Down proteomics: facts and perspectives, Biochem Biophys Res Commun 445 (2014) 683–693.

[52] J.J. Pesavento, Y.B. Kim, G.K. Taylor, N.L. Kelleher, Shotgun annotation of histone modifications: a new approach for streamlined characterization of proteins by top down mass spectrometry, J Am Chem Soc 126 (2004) 3386-3387.

[53] M.V. Holt, T. Wang, N.L. Young NL, High-Throughput Quantitative Top-Down Proteomics: Histone H4, J Am Soc Mass Spectrom 12 (2019) 2548-2560.

[54] R. Amunugama, R. Jones, M. Ford, D. Allen, Bottom-Up Mass Spectrometry-Based Proteomics as an Investigative Analytical Tool for Discovery and Quantification of Proteins in Biological Samples, Adv Wound Care (New Rochelle) 2 (2013) 549–557.

[55] M.S. Lowenthal, Y. Liang, K.W. Phinney, S.E. Stein SE, Quantitative bottom-up Proteomic depends on digestion conditions, Anal Chem 7 (2014) 551-558.

[56] J.R. Yates 3rd, Mass spectrometry and the age of the proteome, J Mass Spectrom 1 (1998) 1-19.

[57] T.D. Lee, J.E. Shively, Enzymatic and chemical digestion of proteins for mass Spectrometry, Methods Enzymol 193 (1990) 361-374.

[58] J.V. Olsen, S.E. Ong, M. Mann, Trypsin cleaves exclusively C-terminal to arginine and lysine residues, Mol Cell Proteomics 6 (2004) 608-614.

[59] A.I. Nesvizhskii, Protein identification by tandem mass spectrometry and sequence database searching, Methods Mol Biol 367 (2007) 87-119.

[60] O. Serang, W. Noble, A review of statistical methods for protein identification using tandem mass spectrometry, Stat Interface 5 (2012) 3–20.

[61] I.V. Chernushevich, A.V. Loboda, B.A. Thomson, An introduction to quadrupole–time-of-flight mass spectrometry, J Mass Spectrom 36 (2001) 849-865.

[62] A.K. Shukla, J.H. Futrell, Tandem mass spectrometry: dissociation of ions by collisional activation, J Mass Spectrom 35 (2000) 1069-1090.

[63] R. A. Yost, C. G. Enke, Selected ion fragmentation with a tandem quadrupole mass spectrometer, J Am Chem Soc 100 (1978) 2274-2275

[64] J. Ponthus, E. Riches, Evaluating the multiple benefits offered by ion mobility-mass spectrometry in oil and petroleum analysis, Int J Ion Mobil Spec 16 (2013) 95–103.

[65] Giles, K. Travelling wave ion mobility, Int J Ion Mobil Spec 16 (2013) 1–3.

[66] J.M. Wells, S.A. McLuckey, Collision-induced dissociation (CID) of peptides and Proteins, Methods Enzymol 402 (2005) 148-185.

[67] V.H. Wysocki, K.A. Resing, Q. Zhang, G. Cheng, Mass spectrometry of peptides And Proteins, Methods 3 (2005) 211-222.

[68] D.C. Stahl, K.M. Swiderek, M.T. Davis, T.D. Lee, Data-controlled automation of liquid chromatography/tandem mass spectrometry analysis of peptide mixtures, J Am Soc Mass Spectrom 7 (1996) 532-540.

[69] L.C. Gillet, A. Leitner, R. Aebersold, Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing, Annu Rev Anal Chem (Palo Alto Calif) 9 (2016) 449-472.

[70] H. Liu, R.G. Sadygov, J.R. Yates 3rd, A model for random sampling and estimation Of relative protein abundance in shotgun proteomics, Anal Chem 76 (2004) 4193-4201.

[71] D. Moran, T. Cross, L.M. Brown, R.M. Colligan, D. Dunbar D, Data-independent acquisition (MSE) with ion mobility provides a systematic method for analysis of a bacteriophage structural proteome, J Virol Methods 195 (2014) 9-17.

[72] J.G. Burniston, J. Connolly, H. Kainulainen, S.L. Britton, L.G. Koch, Label-free profiling of skeletal muscle using high-definition mass spectrometry, Proteomics 14 (2014) 2339–2344.

[73] S. Blech, R. Laux, Resolving the microcosmos of complex samples: UPLC/travelling wave ion mobility separation high resolution mass spectrometry for the analysis of in vivo drug metabolism studies, Int J Ion Mobil Spec 16 (2013) 5–17.

[74] L. Lamont, M. Baumert, N. Ogrinc Potočnik N, Integration of Ion Mobility MS^E after Fully Automated, Online, High-Resolution Liquid Extraction Surface Analysis Micro-Liquid Chromatography, Anal Chem 89 (2017) 11143–11150.

[75] N.J. Bond, P.V. Shliaha, K.S. Lilley, L. Gatto, Improving Qualitative and Quantitative Performance for MSE-based Label-free Proteomics, Journal of Proteome Research 12 (2013) 2340-2353.

[76] P.V. Shliaha, N.J. Bond, L.Gatto, K.S. Lilley, Effects of Traveling Wave Ion Mobility Separation on Data Independent Acquisition in Proteomics Studies, J Proteome Res 12 (2013) 2323-2339.

[77] M. Bantscheff, M. Schirle, G. Sweetman, Quantitative mass spectrometry in proteomics: a critical review, Anal Bioanal Chem 389 (2007) 1017–1031.

[78] P.L. Urban, Quantitative mass spectrometry: an overview, Philos Trans A Math Phys Eng Sci 374 (2016).

[79] V.C. Wasinger, M. Zeng, Y. Yau, Current status and advances in quantitative proteomic mass spectrometry, Int J Proteomics (2013).

[80] A. Doerr, Quantitative mass spectrometry, Nat Methods 6 (2009).

[81] X. Chen, S. Wei, Y. Ji, X. Guo, F. Yang, Quantitative proteomics using SILAC: Principles, applications, and developments, Proteomics 18 (2015) 3175-3192.

[82] K. Kani, Quantitative Proteomics Using SILAC, Methods Mol Biol 1550 (2017) 171-184.

[83] S. Wiese, K.A. Reidegeld, H.E. Meyer, B. Warscheid, Protein labeling by iTRAQ: a New tool for quantitative mass spectrometry in proteome research, Proteomics 3 (2007) 340-350.

[84] L. Zhang, J.E. Elias, Relative Protein Quantification Using Tandem Mass Tag Mass Spectrometry, Methods Mol Biol 1550 (2017) 185-198.

[85] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, B. Kuster, Quantitative mass spectrometry in proteomics: a critical review, Anal Bioanal Chem 389 (2007) 1017-31.

[86] S. Nahnsen, C. Bielow, K. Reinert, O. Kohlbacher, Tools for label-free peptide Quantification, Mol Cell Proteomics 12 (2013) 549-556.

[87] J.C. Silva, M.V. Gorenstein, G.Z. Li, J.P. Vissers, S.J. Geromanos, Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition, Mol Cell Proteomics 5 (2006) 144-156.

CHAPTER 2 – PROTEOMIC PROFILE OF DAPHNIA PULEX USING DATA-INDEPENDENT ACQUISITION MASS SPECTROMETRY AND ION-MOBILITY SEPARATION

2.1 – Introduction

The use of proteomics for functional characterization of model aquatic organisms has emerged as a promising complement to the more traditional genomic approaches. The water flea, *Daphnia pulex*, is a freshwater crustacean that represents a keystone species for the aquatic habitats as well as a useful sentinel species for evidence of environmental stress [1]. *Daphniids* are emerging as model organisms for ecological genomic studies due to their many favourable attributes such as broad geographic and ecological distribution, short generation time, cyclical parthenogenetic reproduction involving the alternation of sexual and asexual reproduction, small genome size, and susceptibility to stress [2]. As a result, many recent studies on *Daphniids* have utilized ecological genomic techniques. While there have been several studies employing various mass spectrometry-based techniques, the *D.pulex* functional proteome remains poorly annotated [1,3-5]. Thus, the continued efforts aimed at characterizing the *D.pulex* proteome in terms of identifyinh novel *D.pulex* proteins and annotating theur putative function via bioiniformatics are required for the development of future target-based assays for monitoring aquatic environmental health, and contributing to the preservation of fresh water resources.

Two major categories of acquisition modes used for proteomic analysis are the so-call data-dependent acquisition (DDA), and data-independent acquisition (DIA) [6]. DDA utilizes an initial survey scan (MS1) prior to isolating a precursor ion for MS/MS analysis

(MS2). While effective at producing high quality fragmentation data for peptide identification, DDA experiments are inherently biased toward higher abundant ions, potentially excluding lower abundant peptide ions [6]. DIA methods attempt to circumvent this issue by simultaneously fragmenting all precursor ions within a specified mass window. The rapid toggling of low and high collision energy scans per duty cycle (MSE) can detect, in parallel, precursor ions and precursor ion fragments, respectively [6]. This enables DIA approaches to better identify lower abundant proteins within complex proteomic studies. DIA methods are enhanced by efforts to reduce sample complexity, such as employing ion mobility (IM) separation techniques. IM is a gas-phase technique that separates ions based on size, shape, and charge [7]. IM has been shown to be effective at separating coeluting peptides and improving proteome coverage during DIA experiments [8]. In this study, we applied for the first time IM-MS and DIA acquisition methods toward the study of the D. pulex proteome

2.2 – Experimental Procedures

Sample Acquisition and Preparation for Mass Spectrometry Analysis

D. pulex samples were collected from Disputed Pond, located in Windsor, Ontario, Canada and maintained as described previously [9]. Briefly, cultures were maintained in Field Laboratory for the Assessment of Multiple Ecological Stressors (FLAMES) medium at 18 ± 1 °C with a circadian cycle consisting of 16 h light:8 h dark within a SHEL LAB biochemical oxygen demand (BOD) incubator [9]. Cultures of D. pulex were allowed to acclimatize to these conditions for a minimum of four generations prior to harvesting. Approximately 50 adult females were placed in algae free media for 12–14 h, filtered using ultra-soft nylon mesh, frozen immediately in liquid nitrogen, and stored at −80 °C. Frozen samples were re-suspended in non-denaturing lysis buffer containing 50 mm Tris-HCl (pH 7.4), 150 mm NaCl, 1% Triton X-100, 0.1% SDS or denaturing lysis buffer containing 8 M urea, 4% CHAPS, 40 mm Tris-HCl (pH7.4), 50 mm DTT (Figure 1A) and supplemented with a HALT (Thermo) protease inhibitor cocktail prior to pulse homogenization. Due to the abundance of gut-based proteases that are resistant to commercially available protease inhibitors and active at lower temperatures, previous attempts to study the D. pulex proteome have demonstrated rapid and extensive protein degradation using standard sample preparation techniques even when homogenized in denaturing buffers (Figure S1, Supporting Information) [4,5,10]. As previously observed by Kemp and Kultz, heating the homogenate at 100 °C for 5 min improves protein degradation due to denaturation of proteases (Figure S1, Supporting Information, left panel) [5]. Homogenate heating using nondenaturing buffers was employed, as denaturation buffers under these conditions have been shown to yield side reactions, such as carbamylation [11]. Preparation of tryptic

peptides for mass spectrometry analysis were generated by both in-gel and in-solution digestion methods to compare the value of pre-fractionation (in-gel) against more rapid and efficient approaches (in-solution). At least two biological replicates were performed for in-gel and in-solution methods. For in-gel digestion methods, SDS/PAGE gels were run using extracts (80 µg) diluted in loading dye (3×) and boiled for 5 min. In-gel digestion proceeded as previously described from 16 gel slices and subsequently pooled to yield 8 fractions at 150 ng µL−1 concentration for MS analysis [12]. For in-solution experiments, 100 µg of protein were precipitated with 13% TCA/acetone for 3 h. Re-suspended pellets were digested overnight at 37 °C with trypsin as described previously and diluted to 150 ng µL−1 for MS analysis [12].

Mass Spectrometry Analysis of Daphnia pulex

Proteomic analysis was performed using a Waters Synapt G2-Si mass spectrometer interfaced with a nanoAcquity ultraperformance liquid chromatography (UPLC) system. Following trypsin digestion, peptides (150 ng µL−1) were injected on a C18 reverse phase column (Waters HSS T3 75 × 150 µm analytical column) at a flow rate of 0.5 µL min−1 using a 3–85% acetonitrile gradient in 0.1% formic acid over 110 min for 1D UPLC analysis. For 2D UPLC separations, peptides (150 ng µL−1) were loaded onto the first dimension XBridge BEH130 C18 nanoEase column (300 × 50 µm), which had been equilibrated with 20 mm ammonium formate, pH 10. First dimension fractions were eluted off using a discontinuous step gradient with acetonitrile elution steps (11%, 14.5%, 17.5%, 20%, 45%, and 65%) at 2 µL min−1 and transferred onto a second trapping column. Peptides were then eluted from the trap onto a nanoAcquity BEH130 C18 analytical column (100 µm × 100 mm, 1.7 µm) using a 0–85% acetonitrile gradient at 400 nL min−1

over 45 min. Eluting peptides were analyzed in DIA mode without IM (MSE) and with IM (HDMSE) by alternating between high and low energy scans in positive ion mode. Precursor ions ranging from 50 to 2000 m/z were scanned at a rate of 0.8 s. A 50 fmol/μL [Glu1]-fibrinopeptide B at 500 nL min−1 was used as a lock mass external calibrant. Cone voltage was set to 30 V and the capillary voltage was kept at 3.8 kV. MSE experiments had fragmentation occurring in the trap region of the TriWave chamber using a linear ramp from 15 to 45 V across the high energy scan. For HDMSE experiments, fragmentation transpired within the transfer region of the TriWave cell following IM separation. Data was collected using MassLynx (version 4.1) and analyzed using Progenesis QI (Nonlinear dynamics version 2.01). Processing parameters for peak detection signal thresholds were 135 counts for low energy (MS1), 30 counts for high energy (MS2), and 750 counts for peptide intensity combined across isotopes and charge states with data lock mass corrected post acquisition. The D. pulex database was downloaded from NCBI on January 15, 2017 containing 33 524 sequences. The protein search parameters were set to include a maximum protein mass of 250 kDa, a minimum of three matched fragment ions per peptide, a minimum of seven fragments ions per protein, and two missed trypsin cleavages was permitted. All single peptide identifications were manually inspected for quality and only accepted if they also met our other criteria (seven fragments per protein, mass accuracy of <10 ppm, and minimum peptide score of 5.0). A 1% false discovery rate was estimated utilizing a decoy database and proteins were required to be detected in at least two technical replicates for positive identification. Variable modifications of carbamidomethylated cysteine and oxidation of methionine were specified. The mass

spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD008455.

2.3 – Results and Discussion

After resolving duplicate identifications and applying our protein hit criteria, a total of 1732 unique *D. pulex* proteins were confidently identified between all methods used **[Figure 2.1C, Table S2.1, Supporting Information]**. Among them, 50% of the proteins were identified in multiple methods with 15% being cross identified between all five techniques. Noticeably, analysis of the results revealed that the use of IM provided the most significant improvement in proteome coverage. This is evident when comparing in-gel and in-solution methods in the absence and presence of IM separation. For both methods, application of IM increased the number of identified peptides/proteins twofold **[Figure 2.1C, Figure S2.2, Supporting Information]**. This effect was not observed when comparing 1D versus 2D UPLC samples. However, the two chromatography approaches did each identify over 100 unique proteins, revealing that some peptides/proteins required 2D UPLC for detection, while other peptides/proteins required the sensitivity advantage associated with 1D separation. The finding that 2D UPLC did not yield a substantial increase in total protein identifications suggests that sample complexity requiring higher peak capacity was not a major factor. This may be a database issue or the result of only partially remedying the activity of Daphnia-specific proteases during sample preparation. Alternatively, ion mobility did result in significant improvement in protein identifications, likely due to a higher degree of orthogonality compared to 2D UPLC as observed by others [8]. Collectively, our results indicate that under our conditions, a comprehensive

chromatography approach employing in-gel and in-solution preparation methods coupled to IM-MS maximizes the proteome coverage of *D. pulex*.

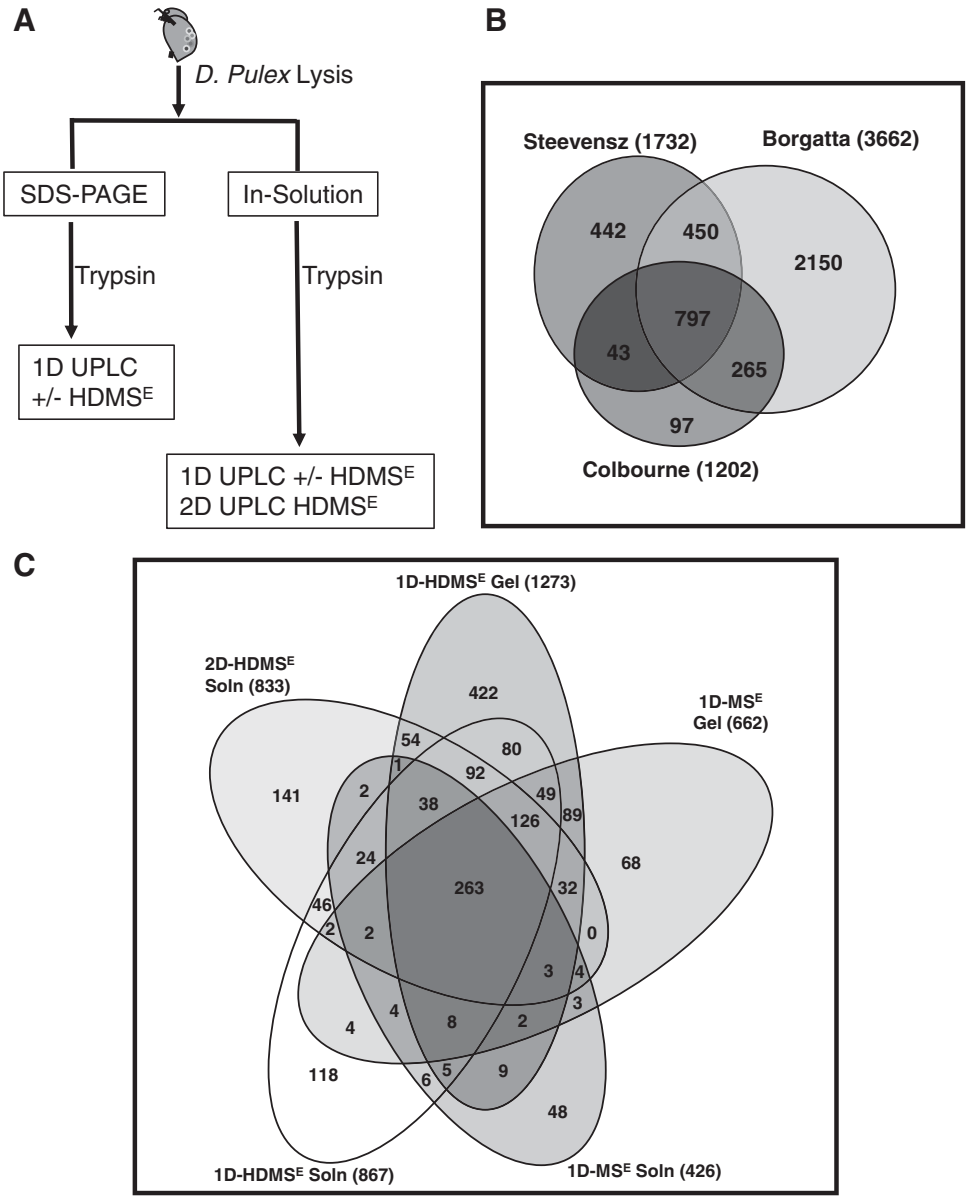


Figure 2.1- Daphnia pulex proteome using DIA-IM-MS. A) Schematic of proteomic workflow. B) Venn diagram comparing proteomic results from three independent investigations including Steevensz et al. (current study), combined runs from Borgatta et al.,[3] and Colbourne et al.[1] C) Venn diagrams comparing proteins identified from proteomic analysis employing in-gel digestion (Gel), in-solution digestion (Soln), data-independent acquisition (MSE), and/or MSE with ion mobility separation (HDMSE)

Due to the fact that the majority of the *D. pulex* proteome remains uncharacterized and/or hypothetical, a large-scale BLAST analysis was performed utilizing the BLAST+ command line user interface.[13] FASTA file sequences of identified proteins were retrieved, along with protein databases containing annotated entries from species of interest (*Homo sapiens*, *Anopheles gambiae*, and *Drosophila melanogaster*), and uploaded into the BLAST+ interface. Batch sequence alignments to identify homologous proteins was performed using an E-value < 0.001 to indicate significant homology. The Panther gene ontology software was used to perform enrichment analysis and ontology determination only on proteins that had a high confident sequence alignment. This resulted in putative annotation of 1427 of the 1732 identified proteins **[Table S2.1, Supporting Information]**. Gene ontology enrichment analysis revealed that while the majority of gene families were equally represented, there were several gene families that appear to be overrepresented in the *D. pulex* proteome compared to the proteomes of closely related invertebrates and humans **[Table S2.2, Supporting Information]**. Most notably, several enzyme families associated with mitochondrial dynamics (dehydrogenases, oxidoreductases), lysosomal function (hydrolases), proteolysis (proteases, metalloproteases), and antioxidant defense (oxidases, reductases) were significantly enriched in our isolated proteome **[Table S2.2, Supporting Information]**. It is noted that the enrichment in proteases correlates to the observed problematic protease activity during *Daphnia* proteome analysis **[Figure S2.1]** [4,5]. In contrast, genes associated with an immune defense system were significantly underrepresented in our proteomic analysis. This observation may be due to the aseptic laboratory conditions employed or the inherent difficulty in isolating immune related transmembrane proteins for proteomic analysis. Alternatively, it may suggest *D. pulex*

utilizes unique defense mechanisms for *preventing pathogenic invasion as supported by previous bioinformatic analysis of the D*. pulex genome [14]. It is important to note that since the *D. pulex* genome is poorly annotated, the gene ontology findings will require re-examination once *D. pulex* genome is better annotated.

We also performed a meta-analysis comparing our data using DIA-IM-MS with the largest two *D. pulex* proteomic studies that employed DDA-MS on orbitrap instruments **[Figure 2.1D]** [1,3,4]. Collectively, a total of 4244 unique *D. pulex* proteins have been identified between the three studies, 442 of which were uniquely identified in our current study **[Figure 2.1D]**. Interestingly, of the 442 unique proteins, 75% were identified only with IM **[Figure 2.2A]**, emphasizing the impact of ion mobility separation for acquiring deeper proteome coverage. Moreover, when we subdivided the IM identified proteins amongst the different sample preparation methods, there was a relatively equal distribution of proteins identified amongst the techniques **[Figure 2.2C]**. However, gene ontology analysis revealed that certain classes of proteins were enriched within individual methods **[Figure 2.2D]**. For example, membrane-associated proteins were found overrepresented in samples prepared using in-solution trypsin digestion compared to in-gel digestion samples with methods employing 2D UPLC identifying the most membrane proteins **[Figure 2.2D]**. This is likely due to inherent deficiencies of gel electrophoresis in processing hydrophobic proteins [15]. For the in-gel digested samples, besides an enrichment of proteolytic enzymes, there was no observable trend in protein class preference with in-gel digestion. Proteins identified exclusively with this sample preparation method likely either benefit from gel electrophoresis-mediated separation or are not amendable to precipitation prior to in-solution digestion.

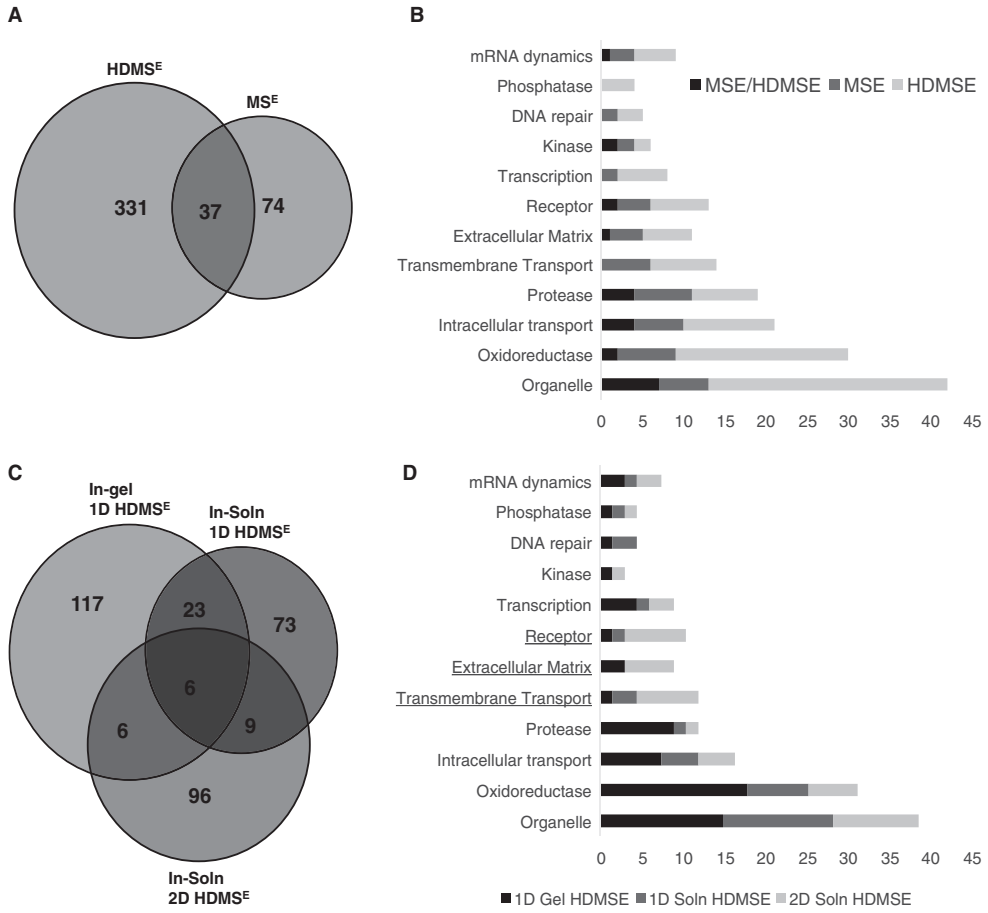


Figure 2.2 - Contribution of ion mobility separation for Daphnia pulex proteomic analysis. A,B) Venn diagram and gene ontology analysis of novel proteins identified in the current study identified without (MS^E) and using ion mobility separation (HDMSE). C,D) Venn diagram and gene ontology analysis comparing the sample preparation techniques used for identifying novel proteins by $HDMS^E$ analysis.

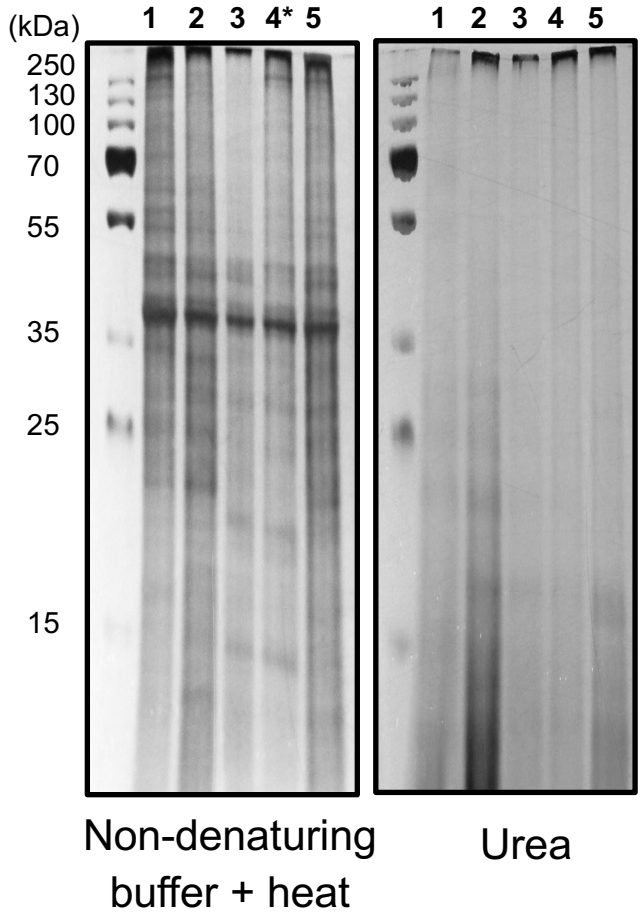
As mentioned above, we were not able to assign a putative description to approximately 300 proteins during our BLAST+ analysis. To obtain some insight into their function, we broadly searched against sequences from all organisms represented in the NCBI non-redundant database. This resulted in numerous proteins predicted to possess various functional domains **[Table 2.1]**. One of the more intriguing findings were the large number of identified proteins predicted to possess domains associated with retrotransposition and other genome modifying mechanisms. This finding provides supporting evidence to the long-term mutation accumulation experiments, which revealed that *Daphnia* has a high rate of large-scale duplications and deletions, with such events often spanning coding regions [16,17]. Surveys on DNA transposons in natural populations of *Daphnia* have also revealed that cyclical parthenogenetic lineages of *Daphnia* exhibit high loads of transposable elements and signatures of recent transpositions [1,18-20]. It will be therefore interesting to examine if these novel putative DNA modifying enzymes are important drivers of DNA transposition in *D. pulex*.

Table 2.1- Annotation of putative domains in Daphnia pulex.

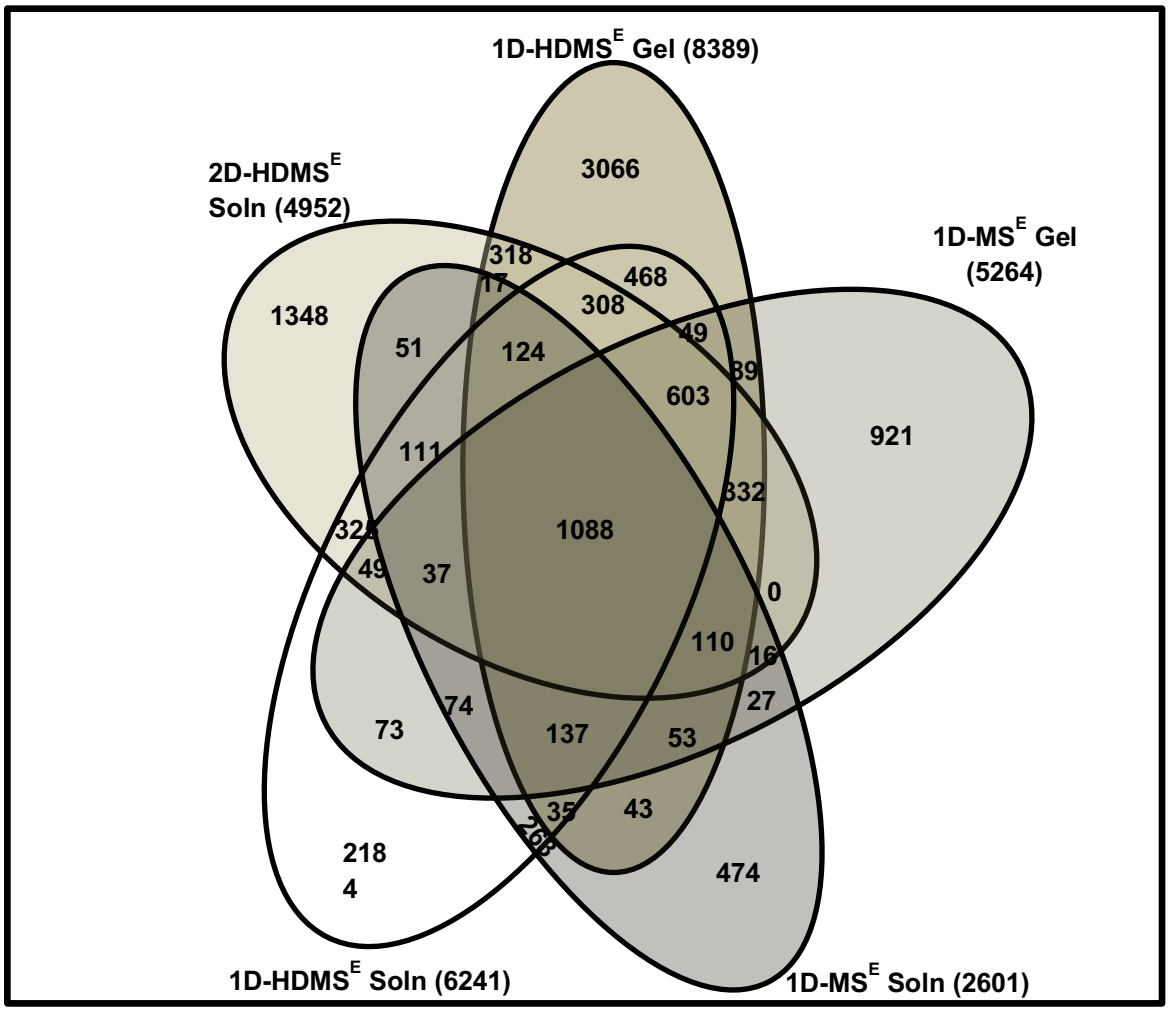
Accession	Domain name	Domain function	E-value
EFX61230.1	Exonuclease/endonuclease/phosphatase	Phosphodiesterase	6.59e-20
EFX71370.1	DDE endonuclease	DNA transposition	2.84e-18
EFX69685.1	Integrase	DNA insertion	4.98e-4
EFX74075.1	DNA Pol3 gamma 3	DNA replication	6.16e-3
EFX62967.1	Structural maintenance of chromosomes	Chromosomal organization and dynamics	5.06e-3
EFX63917.1	Adenylation DNA ligase like	DNA dynamics	1.15e-7
EFX61763.1	Nucleotide binding domain 94	ATP/ADP binding	2.00e-3
EFX62501.1	iSH2_Pi3K	Regulates PI3K activity	3.11e-3
EFX63819.1	Globin like	Oxygen transport	5.10e-25

In conclusion, the current study advances our current knowledge of the *D. pulex* proteome using DIA techniques. We also demonstrate the utility of ion mobility separation in significantly improving proteome coverage and as a valuable tool for future exploration of the *D. pulex* proteome.

2.4 – Supplementary Figures



Supplementary Figure S2.1 - SDS/PAGE analysis of various precipitation procedures. Samples were lysed using non-denaturing buffer and heated at 100 °C for 5 min (Left panel) or Urea buffer (Right panel) and 100 µg of protein was subjected to the following precipitation buffers: (1) No precipitation, (2) Acetone, (3) 13% TCA, (4) 13% TCA/Acetone, and (5) Chloroform/Methanol. The TCA/Acetone buffer was utilized for MS analysis (indicated by an asterisk (*)) as it reproducibly yielded the highest protein yield.



Supplementary Figure S2.2 - Venn diagram comparing peptides identified from proteomic analysis employing in-gel digestion (Gel), in-solution digestion (Soln), data-independent acquisition (MS^E), and/or MS^E with ion mobility separation ($HDMS^E$).

2.5 – References

[1] J. K. Colbourne, M. E. Pfrender, D. Gilbert, et al., The Ecoresponsive Genome of *Daphnia pulex*, Science 331 (2011) 555-561.

[2] B. D. Eads, J. Andrews, J. K. Colbourne, Ecological genomics in Daphnia: stress responses and environmental sex determination, Heredity 100 (2008) 184-190.

[3] M. Borgatta, C. Hernandez, L. A. Decosterd, N. Ch`evre, P. Waridel, Shotgun ecotoxicoproteomics of Daphnia pulex: biochemical effects of the anticancer drug tamoxifen, J Proteome Res 14 (2015) 279-291.

[4] T. Frohlich, G. Arnold, R. Fritsch, T. Mayr, C. Laforsch, LC-MS/MS-based proteome profiling in Daphnia pulex and Daphnia longicephala: the Daphnia pulex genome database as a key for high throughput proteomics in Daphnia, BMC Genomics 10 (2009).

[5] C. J. Kemp, D. Kultz, Controlling proteome degradation in Daphnia pulex, J Exp Zool A Ecol Genet Physiol 317 (2012) 645-651.

[6] L. C. Gillet, A. Leitner, R. Aebersold, Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing, Annu Rev Anal Chem 9 (2016) 449-472.

[7] A. B. Kanu, P. Dwivedi, M. Tam, L. Matz, H. H. Hill, Ion mobility-mass spectrometry, Mass Spectrom 43 (2008) 1-22.

[8] N. J. Bond, P. V. Shliaha, K. S. Lilley, L. Gatto, Improving qualitative and quantitative performance for MS(E)-based label-free proteomics, J Proteome Res 12 (2013) 2340-2353.

[9] P. M. Celis-Salgada, A. Cairns, N. Kim, N. D. Yan, The FLAMES medium: a new, soft-water culture and bioassay medium for Cladocera, Verh Internat Verein Limnol 30 (2008) 265.

[10] R. D¨olling, D. Becker, S. Hawat, M. Koch, A. Schwarzenberger, B. Zeis, Adjustments of serine proteases of Daphnia pulex in response to temperature changes, Comp Biochem Physiol B Biochem Mol Biol 194 (2016) 1-10.

[11] L. Kollipara, R. P. Zahedi, Protein carbamylation: in vivo modification or in vitro artefact?, Proteomics 13 (2013) 941-944.

[12] A. Shevchenko, J. Havli, H. Tomas, J. V. Olsen, M. Mann, In-gel digestion for mass spectrometric characterization of proteins and proteomes, Nat Protoc 1 (2007) 2856-2860.

[13] C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T. L. Madden, BLAST+: architecture and applications, BMC Bioinformatics 10 (2009).

[14] S. J. McTaggart, C. Conlon, J. K. Colbourne, M. L. Blaxter, T. J. Little, The components of the Daphnia pulex immune system as revealed by complete genome sequencing, BMC Genomics 10 (2009).

[15] L. Castellanos-Serra, Y. Ramos, V. Huerta, An in-gel digestion procedure that facilitates the identification of highly hydrophobic proteins by electrospray ionization-mass spectrometry analysis, Proteomics 5 (2005) 2729-2738.

[16] S. Xu, A. Omilian, M. E. Cristescu, High rate of large-scale hemizygous deletions in asexually propagating Daphnia: Implications for the evolution of sex, Mol Biol Evol 28 (2011) 335-342.

[17] J. M. Flynn, F. J. Chain, D. J. Schoen, M. E. Cristescu, Spontaneous mutation accumulation in *Daphnia pulex* in selection free versus competitive environments, Mol Biol Evol 34 (2017) 160-173.

[18] P. Valizadeh, T. J. Crease, The association between breeding system and transposable element dynamics in *Daphnia pulex,* J Mol Evol 66 (2008) 643-654.

[19] S. Schaack, E J. Pritham, A. Wolf, M. Lynch, DNA transposon dynamics in populations of *Daphnia pulex* with and without sex, Proc R Soc B 277 (2010) 2381-2387.

[20] E. H. Penton, T. J. Crease, Evolution of the Transposable Element *Poke*y in the Ribosomal DNA of Species in the Subgenus *Daphnia* (Crustacea: Cladocera), Mol Biol Evol 21 (2004) 1727-1739.

CHAPTER 3 - PROTEOMIC CHARACTERIZATION OF SEMINAL PLASMA FROM ALTERNATIVE REPRODUCTIVE TACTICS OF CHINOOK SALMON (*ONCORHYNCHUS TSWATCHYSHA*)

3.1 - Introduction

In external fertilizing fishes (including salmonids), the efferent duct produces seminal plasma, providing an optimal ionic environment for maintaining the viability of spermatozoa after their release from the testes into the sperm duct and once released into the spawning environment [1]. In addition to maintaining the quality of spermatozoa inside the reproductive tract or in the spawning environment, recent research has shown that seminal plasma also plays a role in sperm competition [2]. Competition occurs when spermatozoa from multiple males contend over the chance to fertilize a female's eggs [2] and is especially prevalent in fish species in which male alternative reproductive tactics are present [3]. The most common alternative reproductive tactic system seen across fish taxa is the sneak-guard dichotomy in males [4]. Sneaker males usually have small body size and use covert techniques to sneak into mating events between guard males and females to obtain reproductive success. In contrast, the guard males are typically large in body size and have more pronounced secondary sexual characteristics to aid in asserting dominance, including fighting off other males while protecting and monopolizing females. The prevailing theory in the field predicts that sneaker males invest more into spermatogenesis instead of body size, resulting in higher quality sperm, while the guard males invest more into body size and secondary sexual characteristics in order to defend females or resources important to females [5]. For example, in Atlantic salmon (*Salmo salar*), the precocious parr (sneaker male), relative to body size, has larger testes, ejaculate volume, number of

spermatozoa cells, higher spermatozoa motility, and live longer compared to the anadromous (guard) males [6,7].

Chinook salmon (*Oncorhynchus tshawytscha*) also exhibit the sneakguard alternative reproductive tactic in males, where the large, dominant hooknose (guards) have priority in mating positions with females, while the small, precocious jack males (sneakers) adopt the sneaking tactic [8,9]. Previous work has shown that jacks have relatively larger testes and their spermatozoa swim faster in river water compared to hooknose [10], supporting the theory that sneaker males (jack) possess higher sperm quality in general [11]. However, differences in testes size may also result in seminal plasma variation between jack and hooknose males that could impact fertilization success.

Most of the studies to date that examine sperm competition dynamics in fishes have focused on either looking at differences in spermatozoa number or spermatozoa quality [12]. However, spermatozoa only make up a portion of the ejaculate and other components, such as seminal plasma (fluid) can have effects on the outcome of sperm competition [13,14]. For example, Locatello and colleagues [15] demonstrated that for the grass goby (*Zosterisessor ophiocephalus*) the seminal plasma from the sneaker tactic can attenuate the performance of the rival guards' spermatozoa. Furthermore, spermatozoa from sneaker males were shown to achieve higher velocity and fertilization success by utilizing the guard male's seminal plasma [15].

Recent advances in quantitative proteomics have given rise to labelfree techniques for performing comparative analysis of protein levels between different samples [16].

Label-free techniques typically compare normalized precursor ion intensities across multiple biological and technical replicates to detect differential protein levels [16]. Furthermore, label-free approaches combined with data independent acquisition modes and orthogonal separation techniques, such as ion mobility separation, greatly enhances reproducibility, sensitivity, and sequence coverage compared to data dependent spectral counting techniques [17,18]. Thus, the promise of proteomics has made it a powerful approach to analyze protein dynamics in the male gametes of aquatic animals [19–22]. However, despite the increasing evidence that seminal plasma is an important factor influencing spermatozoa quality and fertilization competition, little information is available on its protein composition in Chinook salmon. In this study, we aimed to advance the understanding of the Chinook salmon seminal plasma proteome and determine functional differences between alternative reproductive tactics of Chinook salmon from a wild spawning population using label-free quantitative proteomics.

3.2 – MATERIALS AND METHODS

Collection of Chinook Male Tactics

Sexually mature Chinook salmon were collected during two spawning sessions (2–6 October, 2013 and 4–6 October, 2014) using backpack electrofishing from a winter run in the Credit River (Mississauga, Ontario, Canada, N 43°35′, W 79°42′), which flows into Lake Ontario (that has been stocked for over 40 years) [11,23–25]. We considered males b675 mm in fork length to be jacks based on lengthage relationships derived from otolith assessment for our population. All males N675 mm were considered hooknose males and documented as age 1+ (675–750 mm), age 2+ (750–850 mm) or age 3+ (900– 950 mm) [8,10–11] (Ontario Ministry of Natural Resources, unpublished data). In addition, all fish b675 mm in fork length in our population were morphologically characteristic of jacks (e.g. coloration similar to that of females).

Chinook salmon were located upstream in turbid water ranging from 2 to 4 ft in depth. Water temperature at the time of collection was ~11 °C. The Ontario Ministry of Natural Resource and Forestry technicians (License number 1081183) humanely sacrificed the fish and obtained milt (fluid containing spermatozoa and seminal plasma) samples by applying pressure on the abdomen of each fish. The initial male ejaculate was discarded in a standardized manner and the external urogenital pore was wiped dry to avoid contamination from water, urine, feces, and blood. Milt samples were collected into 532 mL clear Whirl-Pak sample cooling bags (Nasco, Newmarket, ON, Canada) and were stored on ice in coolers for approximately 4 h during transport back to the laboratory for immediate processing. All research followed the University of Windsor and Canadian Council of Animal Care guidelines (AUPP 14–25).

In-Solution Trypsin Digest of Seminal Plasma Samples

The milt samples from five biological replicates per tactic (5 jack males and 5 hooknose males) were subjected to centrifugation at (300 ×g) for 10 min to separate seminal plasma and spermatozoa cells. Seminal plasma protein concentration was determined by Bradford assay. Equal amounts of total protein from each biological specimen was precipitated overnight using a chilled 4:1 acetone solution at −20 °C. Protein pellets were washed three times with acetone and resuspended in solubilization buffer containing 0.1% RapiGest SF surfactant (Waters) in 50 mM ammonium bicarbonate (Fisher Scientific). Protein samples were reduced and alkylated with 10 mM dithiothreitol (Fisher Scientific) and 55 mM iodoacetamide (Sigma), subjected to insolution trypsin digestion (Promega) at a 1:50 (enzyme:protein) ratio and incubated overnight at 37 °C. Trifluoroacetic acid (Pierce) was added to each sample the next day at a final concentration of 0.5% to quench enzyme activity and hydrolyze the surfactant according to manufacturer's protocol. All samples were de-salted using OASIS HLB extraction columns (Waters). Peptides were eluted in 65% acetonitrile (Burdick & Jackson) and concentrated by vacuum centrifugation. The peptides were then re-suspended in 0.1% formic acid, 3% acetonitrile and supplemented with Hi3 internal peptide standards (Waters) at a final concentration of 12.5 fmol/µL for quantitation.

Liquid Chromatography-Mass Spectrometry

Generated tryptic peptides were separated on a 1.7 µm BEH130 100 µm × 100 mm reverse phase column (Waters) at a flow rate of 0.3 µL/min using a nanoAcquity UPLC system (Waters). Mobile phase A consisted of 0.1% formic acid in H2O and mobile phase

B consisted of acetonitrile with 0.1% formic acid. From a 5 µL sample loop, 4 µL was injected onto a 180 µm × 20 mm C18 pre-column trap for 3 min at 5.0 µL/min. Peptides were then separated on a 120 min gradient (3– 25% B for 70 min, 25–50% B for 20 min, 50–85% B for 9 min, 85–3% for 2 min, 3% B for 20 min) and electrosprayed into a SYNAPT G2-Si mass spectrometer (Waters). The mass spectrometer was operated in data independent acquisition mode employing ion mobility separation (HDMSE) alternating between low energy (4 eV) and high energy (20–45 eV) scans in positive resolution mode scanning from 50 to 2000 m/z with a scan rate of 0.6 s. [Glu1]-fibrinopeptide B (50 fmol/µL) was used as the lock mass external calibrant. Cone voltage was set to 30 V, and the capillary voltage was set at 3 kV. MassLynx (version 4.1) was used to collect the data. The raw data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD005533.

Processing was performed using Progenesis QI (Nonlinear Dynamics) for peptide identification analysis and a teleost subset of the NCBI database (the teleost database consisted of 1,000,974 proteins and was downloaded on October 10, 2014). The following processing parameters were used: low energy noise reduction thresholds-135, high-energy threshold-30, and intensity threshold-750, with data lock mass corrected post acquisition. Search parameters included a maximum protein mass of 250 kDa, a minimum of two matched peptides per protein, a minimum of seven fragment ion matches per protein, a minimum of three fragment ion matches per peptide, a maximum of two missed cleavages (trypsin). The false-discovery rate was estimated to be below 1% using a decoy database. Variable modifications for carbamidomethylated cysteine and oxidation of methionine were specified. Three technical replicates for each of the five biological replicates were

analyzed per tactic (30 total measurements). Quantitation was determined using the three most abundant peptides per protein (including the internal standard) and normalized using the Progenesis QI software. The normalization processing utilized a reference chromatography run and spiked Hi3 internal standard peptides (50 fmol on column) to convert raw abundance values to absolute measurements (fmol on column) [26].

Statistical Analysis and Bioinformatics

Differences in protein abundance was determined by excluding proteins identified with less than three matched peptide products and with an ANOVA score (p-value) < 0.05. A Student's t-test was performed comparing the average amount of protein (fmol) from the five biological replicates of hooknose and jack males. Protein values for each biological replicate were obtained by calculating the average protein abundance from each corresponding technical replicate (in triplicate). The p-values obtained were corrected for false-discovery by applying the BenjaminiHochberg method of analysis to produce a corrected t-test value (q value) [27]. All q-values < 0.05 were considered significant and protein abundance ratios >1.5 were considered significant. A Volcano-plot was constructed by using the “ggplot2” package on “R” statistical software by plotting the log2fold change (x-axis) in relation to the −Log10 p-value (y-axis) [28,29]. Gene ontology was assigned to the identified seminal plasma proteins by sequence alignment using BLAST+ (Version 2.2.30) against the Human Uniprot/SwissProt database. The protein accession numbers obtained from the BLAST+ analysis were uploaded onto UniprotKB to map gene ontology.

Immunoblot Analysis

For immunoblots assays, 10 µg of seminal plasma protein from selected biological specimens were loaded onto 12% SDS/PAGE gels. Alternatively, 25 µg of pooled seminal plasma from jack and hooknose samples was subjected to exosome isolation using the Total Exosome Isolation Kit (Invitrogen) according to the manufacturer's instructions. Following SDS/PAGE analysis, the samples were electrotransferred onto PVDF membranes and incubated for 1 h at room temperature with a cross reacting rabbit lactate dehydrogenase (LDH) antibody raised against LDH from *Toxoplasma gondi* (kind gift from Dr. Sirinart Ananvoranich) at a 1:5000 dilution in 2.5% dry milk (v/v)/TBST (20 mM Tris Base, 140 mM NaCl, 0.1% Tween 20, pH = 7.6) or a cross reacting rabbit pan 14-3-3 antibody raised against human 14-3-3 (Santa Cruz Biolabs) at a 1:5000 dilution in 2.5% dry milk/TBST and anti-rabbit HRP (Biorad) at a 1:5000 dilution made with 2.5% milk/ TBST at room temperature for 45 min. Proteins were visualized using Super Signal West Femto Reagent (Thermo Scientific).

3.3 – RESULTS AND DISCUSSION

Identification of Hooknose and Jack Seminal Plasma Proteins by Data Independent Ion Mobility Mass Spectrometry

To better characterize the seminal plasma protein profiles of hooknose and jack Chinook salmon males, a cross-species proteomics and bioinformatics workflow was utilized **[Figure 3.1]**. Ten sexually mature biological specimens (five jack males and five hooknose males) were collected from the field and the milt was isolated. It is important to note that jack males sexually mature one year earlier than hooknose males [8–11]. Moreover, during fertilization competition, a younger jack male will generally be competing with an older hooknose male (younger in terms of age, but equal in terms of sexual maturity). Thus, although we cannot rule out contributions of age to differences in the seminal plasma proteome, it was imperative to compare the seminal plasma proteome at the stage of equal sexual maturity for the purpose of investigating protein factors contributing to fertilization competition.

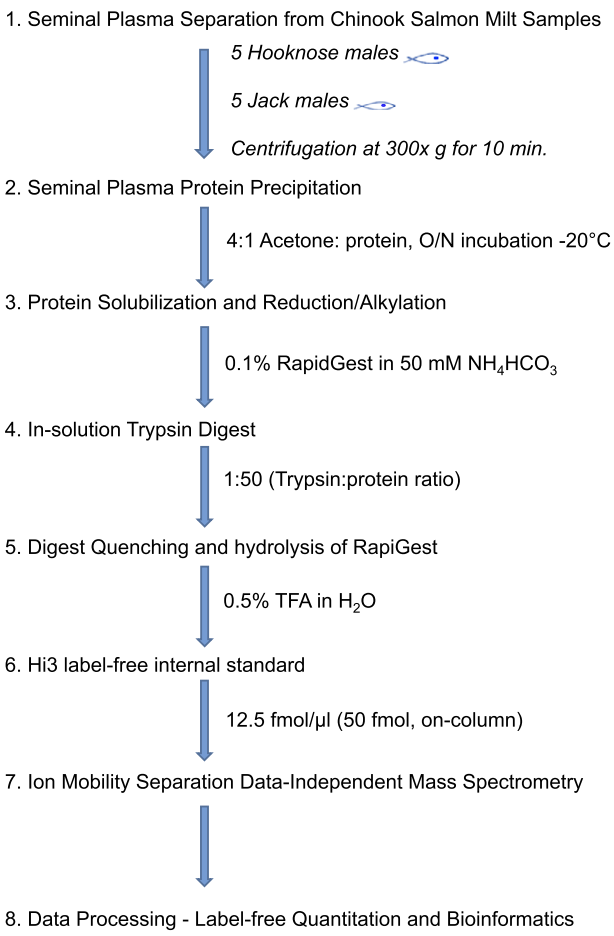


Figure 3.1- Schematic workflow of seminal plasma sample preparation and proteomic analysis. Seminal plasma was extracted from milt by centrifugation, and proteins were prepared for mass spectrometry analysis using RapiGest solubilization and in-solution trypsin digestion (Steps 1–5). Label-free internal standards (Hi3) were added to each sample for absolute quantitation (Step 6). Samples were analyzed by UPLC ion mobility data-independent mass spectrometry (Step 7) and the data was processed using Progenesis-QI (Step 8). Statistical analysis was performed to determine significant differences in protein abundance (Step 8).

Seminal plasma was isolated and processed for mass spectrometry analysis (see "Materials and Methods" for details). Furthermore, synthetically prepared internal standards based on peptide sequences found in the *E. coli* protein ClpB (Hi3) were added to all biological replicates for absolute quantitative analysis. Protein identification was accomplished using data independent tandem mass spectrometry (LCMSE). To increase the number of unique proteins identified and increase protein sequence coverage, ion mobility separation of the peptides was performed using the traveling wave function of the SYNAPT G2-Si system located between the quadrupole and time of flight cell. Preliminary scout runs determined that incorporating ion mobility separation in the workflow increased the number of proteins identified by ~40% (data not shown) demonstrating the significant improvement in proteome coverage when ion mobility separation is employed as reported by other proteomic studies employing ion mobility separation [18,30]. In total, three technical replicates were analyzed for each of the ten biological replicates (five from hooknose and five from jack males).

The Chinook salmon genome is poorly annotated [31] with available protein databases having <900 putative protein sequences for proteomic analysis. With the knowledge that conservation between closely related organisms increases the probability of identification in poorly characterized species [32–34], the *teleost* (bony fish) database was utilized for protein identification. Whole sequence protein alignments resolved conflicts and filtered out protein redundancy, while BLAST analysis assigned protein identification to predicted proteins **[Table S3.1]**. After eliminating proteins identified by one or two peptides and proteins not detected in all thirty technical replicates, our analysis yielded 345 total proteins found in Chinook salmon seminal plasma **[Tables S3.1 and**

S3.2]. Furthermore, the addition of Hi3 internal standards to all of the biological replicates allowed for label-free quantitation of identified proteins using the Progenesis QI software package [16,35]. Analysis of relative abundance and calculated on-column peptide levels (fmol) revealed a dynamic range that spanned 4 orders of magnitude, although the majority of proteins were within a dynamic range of 3 orders of magnitude (0.1–500 fmol on column) **[Figure 3.2]**. At the high end of the dynamic range were well-characterized seminal plasma proteins such as transferrin, albumin-2, and alkaline phosphatase while at the low end of the dynamic range were known seminal plasma proteins DJ-1 and ceruloplasmin [36]. Collectively, LC-MSE using ion mobility separation yielded the first described proteomic data set for Chinook salmon seminal plasma, and to our knowledge, the largest seminal plasma proteome reported for teleost fish.

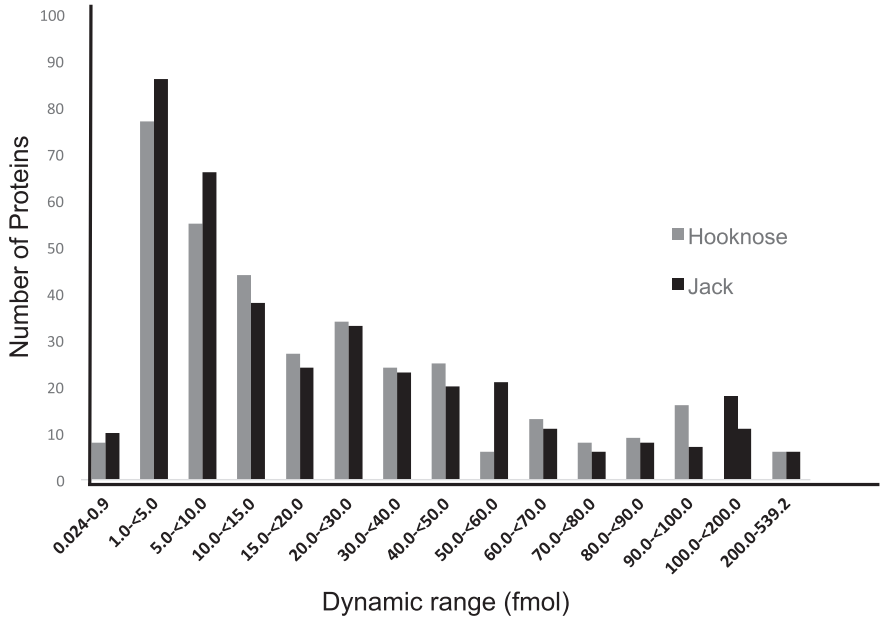


Figure 3.2 - The dynamic range of on-column protein abundance identified from Hooknose and Jack seminal plasma. The distribution of proteins (y-axis) within a range of on-column protein abundances in fmol (x-axis) for both hooknose and jack males. Grey bars represent hooknose proteins and black bars represent jack proteins.

Seminal Plasma Proteins Identified in Hooknose and Jack Chinook Salmon Males

Gene Ontology (GO) bioinformatic analysis provides an overview of the dominant functional traits of isolated proteomes. Of the 345 proteins identified that were common in the seminal plasma fluid of both Chinook male species, 269 (78%) were successfully mapped to a GO **[Figure 3.3]**. For the biological process category, metabolism and cellular regulation/signaling are most represented in the proteome **[Figure 3.3A]**. With regards to molecular function, the predominant gene function activities represented in the Chinook seminal fluid are binding activity and catalytic activity **[Figure 3.3B]**. A similar GO profile was observed in the closely related rainbow trout seminal plasma [37] and carp seminal plasma proteome studies [38], providing a level of validation for the identified proteins. GO analysis of our findings further supports the notion that the principle components of teleost seminal fluid consists of immunomodulators and redox regulators for spermatozoa survival, metabolic factors for energy requirements during fertilization, and proteolytic enzymes/inhibitors for temporal regulation of these processes **[Figure 3.4**, **Table S3.1]**.

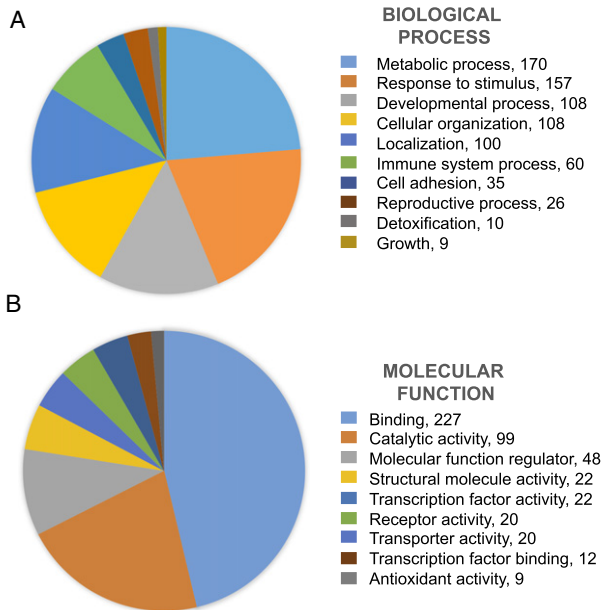


Figure 3.3 - Gene ontology of seminal plasma proteins identified in Chinook salmon (Oncorhynchus tswatchysha) jack and hooknose seminal plasma. A) Gene ontology mapped for seminal plasma proteins in relation to biological process. B) Gene ontology mapped for seminal plasma proteins in relation to molecular function. Gene ontology terms are shown in adjacent legend with corresponding number of matching proteins

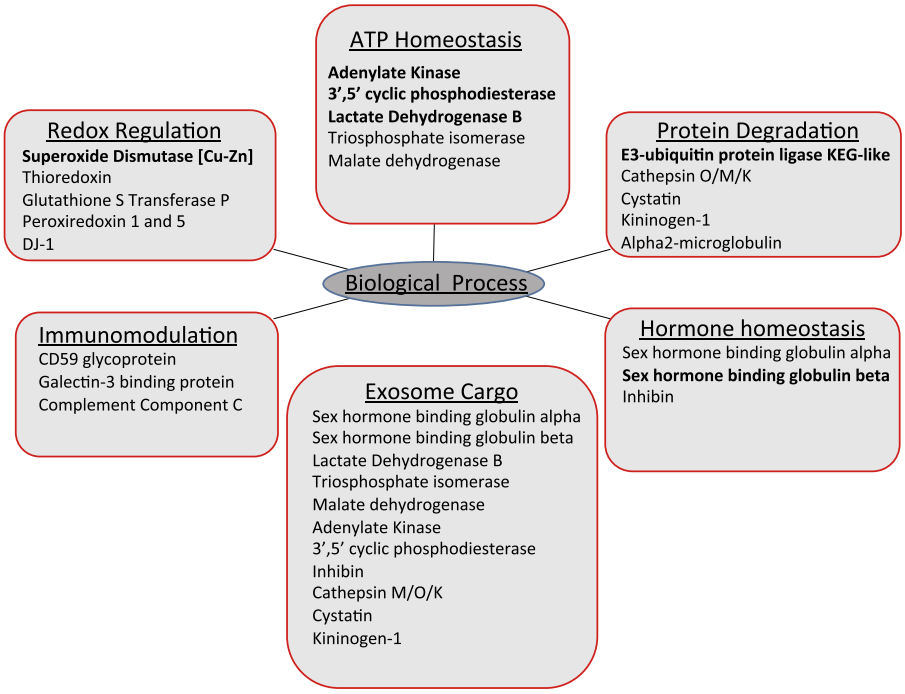


Figure 3.4 - Representative seminal plasma proteins identified in Chinook salmon. Identified proteins of interest were categorized under biological processes known to be critical for seminal plasma biology. Highlighted in bold are proteins whose protein abundances were observed to be statistically significant (**Fig. 3.7**).

In addition to a common gene ontology profile, most of the specific proteins identified in previous fish seminal plasma proteomic studies from rainbow trout [37] and carp [38] were also identified in our dataset **[Table S3.1]**. For instance, several members of the complement component family (C3, C3–3, C3–4, C6, C8α, C8β and C9) were identified. These proteins have a well-characterized role in the innate immune response [39] and may possibly be associated with sperm competition [39,40]. Apolipoproteins (AI, AII, AIV, E and F) involved in lipid transport, metabolism, as well as the immune response [41] were also identified in our study, consistent with past fish seminal plasma proteomic studies [37,38].

Potential Chinook Salmon Biomarkers Relevant to Sperm Fitness

In addition to the complement of proteins reported in previous teleost studies, the present proteomic work also identified novel teleost seminal plasma proteins, including certain proteins that may be suited to serve as biomarkers related to sperm quality **[Table S3.1]**. One example is the identification of inhibin, which is novel for fish seminal fluid, but has been identified in human seminal plasma [36]. Inhibin is an endocrine factor that appears to regulate hormone homeostasis for female and male reproductive organs and has been proposed to be a biomarker of spermatogenesis [42,43]. Another example is the protease inhibitor alpha-2-microglobulin (α2M) [44,45], discovered in the seminal plasma of both tactics **[Figure 3.4**, and **Table S3.1]**. α2M has been proposed to serve as a biomarker for human sperm quality as increasing concentrations of α2M have been correlated to progressive forward motility [46]. α2M has a broad specificity of protease targets and thus affects a diverse range of cellular processes [47]. Interestingly, the aforementioned protein inhibin, is a known binding partner of α2M [48]. The levels and

activities of inhibin and α2M in Chinook salmon warrant further exploration to determine their suitability as a biomarker for spermatozoa fitness.

Orthologues of many of the proteins identified in our study have been associated with extracellular vesicles called exosomes **[Figure 3.4]**. Most cell types secrete exosomes and are found in a variety of bodily fluids including seminal plasma [49]. Exosomes are emerging as powerful biomarkers for use in diagnostic and prognostic analysis of a variety of human disorders [49–54]. With regards to seminal plasma, exosomes have been shown to play an immunosuppressive role in protecting spermatozoa from the male immune system [53,54]. Furthermore, seminal exosomes are thought to deliver protein cargo and other biomolecules to the spermatozoa to facilitate fertilization [55–57]. In our data set, both tactics possessed proteins implicated in regulating the export of exosomes, such as rab11 family interacting protein [49,58], and a variety of proteins previously described as exosome cargo **[Figure 3.4]**. Most notably, immunosuppressive proteins CD59, and galectin-3 binding protein were identified and have been shown in other species to reside in exosomes [59–61]. The levels of these and other exosome cargo may be useful indicators of sperm fertilization fitness for both Chinook male tactics.

Quantitative Protein Differences in The Seminal Plasma of Individual Tactics of Chinook Salmon

A major objective of our study was to elucidate differences in protein levels between hooknose and jack seminal plasma as a means to better understand tactic specific reproductive fitness. Proteins identified in all technical and biological replicates were considered for label free quantitative analysis and statistically significant protein differences were determined (see "Materials and Methods" for details). A volcano plot displaying normalized log2 protein abundance ratio against statistical significance measurements (−log10 q value) is shown in **[Figure 3.5.]** Overall, we observed 29 proteins with statistically significant protein level differences and calculated abundance ratios of at least 1.5 **[Figure 3.5** and **Table S3.1]** [62,63]. GO analysis of these altered proteins revealed a similar GO distribution between the tactics and did not deviate significantly from the predominant molecular functions and biological processes of the proteins that are common between hooknose and jack males **[Figure 3.6]**. This suggests that the seminal plasmas of hooknose and jack males generally are functionally equivalent in terms of protein classifications. However, attention to the specific protein differences suggests mechanistic and functional diversity between the tactics, including proteins regulating hormonal transport, ATP metabolism, redox regulation, and proteolysis **[Figure 3.7**, and **Table 3.1]**. For example, sex hormone binding globulin (SHBG) was found at significantly higher levels in jack males than in hooknose males **[Figure 3.7**, and **Table 3.1]**. SHBG is a well-characterized steroid hormone transporter that regulates androgen/estrogen availability to target tissues [64]. Furthermore, the steroid hormone targets for SHBG are known modulators for spermatogenesis, spermatozoa maturation, and motility acting through non-genomic mechanisms [65, 66]. Therefore, higher amounts of SHBG in jack

males suggest a greater capacity for hormone-mediated effects in the seminal plasma of this tactic.

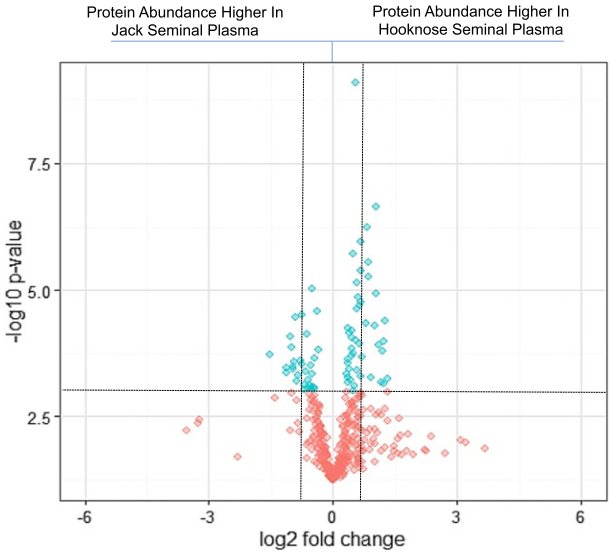


Figure 3.5 - Volcano plot depicting differentially abundant seminal plasma proteins between hooknose and jack males. The volcano plot shows $-\log_{10}$ corrected t-test values (q- value) plotted against $\log_2$ fold change (x-axis). Data points above horizontal dashed line (q values below 0.05) and to the left and right of the vertical dashed lines (abundance ratios above 1.5) represent proteins with statistically significant abundance differences (highlighted in blue).

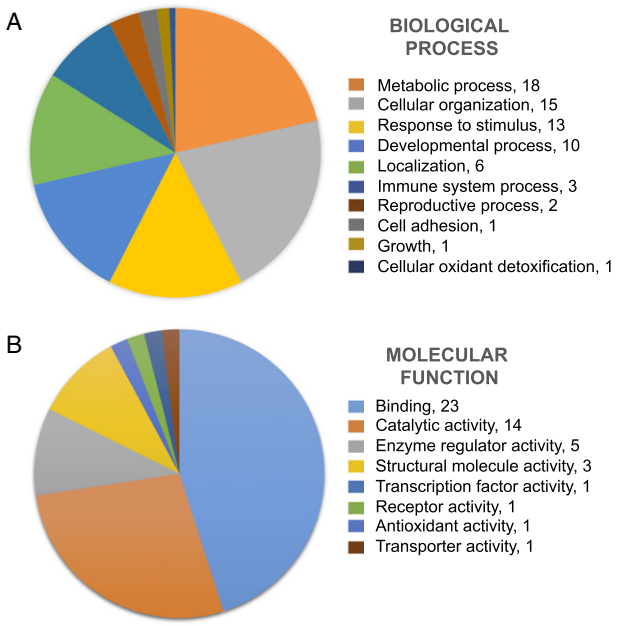


Figure 3.6 - Gene Ontology for seminal plasma proteins differentially abundant in Chinook salmon (Oncorhynchus tswatchysha) jack and hooknose seminal plasma. Gene ontology for proteins whose levels were determined to be statistically significant in jack and hooknose seminal plasma in relation to molecular function (A) or biological process (B)

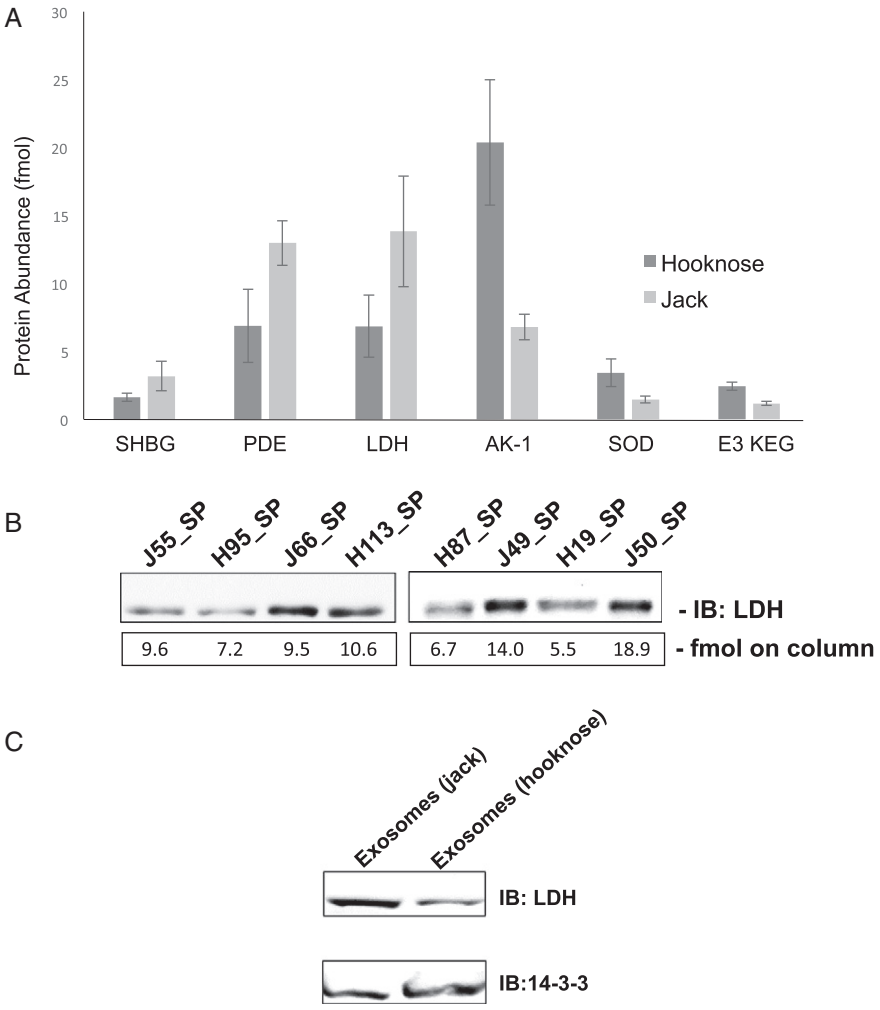


Figure 3.7- Absolute quantitation of selected proteins from jack and hooknose seminal plasma. A) Protein abundances are shown as mean fmol (on-column) ± standard deviation (see Materials and Methods for details). All protein comparisons shown were statistically significant, $p < 0.001$. Individual statistical values are displayed in **Table 3.1**. Proteins shown are the following: Sex hormone binding globulin (SHBG), 3″,5″-cyclic phosphodiesterase (PDE), lactate dehydrogenase (LDH), adenylate kinase 1-1 (AK-1), superoxide dismutase (SOD), and E3 ubiquitin ligase KEG-like (E3 KEG-like). B) Seminal plasma (10 μg) from individual biological specimens was analyzed by LDH immunoblotting using a cross reacting rabbit LDH antibody. The corresponding on-column fmol amount determined by mass spectrometry for the individual biological specimens is shown. C) Isolated exosomes from 25 μg of pooled jack and hooknose seminal plasma were subjected to immunoblot analysis using anti-LDH and anti-14-3-3 antibodies

Table 3.1 - Functional classification of differentially abundant proteins between hooknose and jack males. Proteins whose abundance ratios >1.5 and had p-values and FDR adjusted q-values < 0.05 were considered significant. Proteins are organized by abundance ratios and which tactic contained the higher abundance; hooknose (upper panel), jack (lower panel)

Tactic	*p*-Value	q-Value	% CV	Fold change	Protein description	Protein function
Hook	1.44E-09	0.006025	1.74	2.4	Superoxide dismutase_P03946.2	Redox homeostasis
Hook	2.12E-06	0.041019	2.80	2.3	Adenylate kinase 1_ACH70899.1	ATP shuttling
Hook	2.38E-08	0.019191	4.79	2.2	Spidroin-1-like_XP_014065990.1	Extracellular matrix organization
Hook	1.23E-09	0.015444	2.12	2.1	RNA-binding motif protein_XP_014067707.1	Alternative splicing
Hook	1.67E-15	0.000319	3.76	2.1	E3 ubiquitin ligase KEG_XP_014062379.1	Protein ubiquitination/degradation
Hook	1.90E-05	0.034955	2.08	1.9	Creatine kinase B-type_EMP32891.1	ATP homeostasis
Hook	3.52E-11	0.001165	2.86	1.8	Apolipoprotein C-I_NP_001134834.1	Membrane stabilization
Hook	2.37E-11	0.001747	1.95	1.8	Lumican precursor_ACI33424.1	Extracellular matrix organization
Hook	9.36E-07	0.007213	1.06	1.7	Semaphorin-4F-like_XP_014065278.1	Regulation of cell migration
Hook	7.40E-10	0.000674	4.38	1.6	Transcription factor Sox-19a_XP_014051352.1	Transcription factor
Hook	1.75E-10	0.001586	1.65	1.6	Creatine kinase_P24722.1	ATP homeostasis
Hook	2.83E-06	0.026153	2.84	1.5	Proteasome subunit beta type 5_ACH85299.1	Protein degradation
Jack	4.20E-07	0.023216	11.82	2.9	Actin gamma_NP_001017750.1	Cell motility
Jack	4.11E-05	0.027859	4.22	2.2	Precerebellin protein_AAF04305.2	Inflammatory response
Jack	5.76E-07	0.01682	8.73	2.0	L-lactate dehydrogenase B_XP_014063047.1	ATP production
Jack	1.02E-07	0.010957	6.93	2.0	ARF GTPase protein GIT2_XP_013993786.1	Membrane re-modelling
Jack	3.48E-07	0.026799	3.41	2.0	Sex hormone-binding globulin β_ACJ25982.1	Hormone transport
Jack	2.82E-07	0.006442	2.93	1.9	3′,5′-Cyclic phosphodieserase_XP_014008493.1	cAMP hydrolysis
Jack	2.50E-05	0.043128	4.87	1.8	Ubiquitin hydrolase 44_XP_014063753.1	Protein de-ubiquitination
Jack	9.01E-05	0.036913	3.17	1.8	Creating kinase-2_ACH70915.1	ATP homeostasis
Jack	7.27E-09	0.005507	4.79	1.7	Proteasome subunit alpha type 4_AAH45970.1	Protein degradation

Another intriguing finding was the observed increase in the enzyme cGMP-inhibited 3′,5′-cyclic phosphodiesterase (PDE) in the jack males. PDE catalyzes the hydrolysis of cAMP, a potent second messenger that regulates numerous spermatozoa related activities such as motility and capacitation [67,68]. Extracellular PDE is widely reported in literature and has also been characterized in the seminal plasma of numerous organisms [69,70]. The difference in PDE levels between the tactics insinuates that hooknose males, by virtue of having lower levels of PDE, may exhibit more prominent cAMP signaling than jack males. In addition to factors affecting cAMP activities, levels of enzymes regulating additional metabolic compounds were found to be different between hooknose and jack males. Interestingly, the nucleotide converting enzyme adenylate kinase (AK) was detected at a two-fold higher level in hooknose males **[Figure 3.7, and Table 3.1]**. AK catalyzes the reversible reaction of ATP + AMP = 2ADP and thus represents an important mediator of bioenergetic sensing [71]. AK is detected in extracellular fluids, including seminal plasma [72,73], and has been reported to mediate ATP shuttling to ATPases residing in the spermatozoon tail, facilitating motility [74].

One of the most prominent protein differences was a two-fold abundance ratio increase in lactate dehydrogenase (LDH) in jack males **[Figure 3.7A, and Table 3.1]**. Interestingly, we were able to validate the quantitative proteomic data by immunoblot analysis using a cross reacting LDH antibody **[Figure 3.7B]**. Overall, we observed higher amounts of LDH in jack males compared to hooknose males, although one hooknose individual (H113) displayed elevated LDH levels by immunoblot analysis and quantitative proteomics. Furthermore, as LDH has been characterized to be a component in seminal plasma exosomes [56], we isolated exosomes from pooled seminal plasma and observed

higher levels of LDH from jack males compared to hooknose males **[Figure 3.7C]**. As a control we also probed for 14-3-3 using a pan 14-3-3 antibody. This protein is a known exosome protein [51–52] and was found to be equally abundant in our proteomic analysis as well as in our isolated exosome samples **[Figure 3.7C]**. It has been proposed in several species that spermatozoa can utilize lactate found in the seminal plasma as an energy source, whereby LDH in the seminal plasma converts lactate to pyruvate and NADH that is subsequently transported into the sperm's mitochondria for oxidative phosphorylation and ATP production [75]. This metabolic system is critical for meeting the energy demands of motile spermatozoa and thus is a contributing factor for producing high quality sperm [76]. Furthermore, this finding is consistent with past studies that have predicted jack males to participate in sperm competition by the so-called “loaded raffle” mechanism, whereby sneaker males (jacks) possess higher quality spermatozoa in order to compete for reproductive success [10]. Thus, further detailed mechanistic studies will be important to examine if jack males utilize higher levels of LDH in their seminal plasma to provide their spermatozoa a metabolic advantage over hooknose spermatozoa.

A notable class of proteins identified in the Chinook salmon seminal plasma of both tactics was those that participate in oxidative stress resistance such as thioredoxin, peroxiredoxin, glutathione transferase, and DJ-1 **[Table S3.1]**. Oxidative insults from reactive oxygen species (ROS) during the shift from hypoxic to normoxic conditions during external fertilization can be detrimental to spermatozoa cells [77]. Thus, up-regulation of antioxidant factors enhances sperm viability by increasing survival towards oxygen radicals. ROS have also been shown to have a degradative effect on spermatozoa performance such as motility but help enhance immune function and increase capacitation

in humans [78]. Strikingly, our study found superoxide dismutase (Zn/Cu SOD) levels to be greater than two-fold higher in hooknose males **[Figure 3.7A, and Table 3.1]**. SOD levels in seminal plasma have been shown to correlate to ROS tolerance, protecting spermatozoa from oxidative damage [79]. However, recently reports have demonstrated that SOD activity inhibits spermatozoa capacitation by eliminating localized superoxide anions participating in spermatozoa membrane remodeling [80]. Future studies will be required to determine the significance of its elevated levels in hooknose males as it relates to sperm competition with jack males.

The importance of proteolysis within seminal plasma and sperm maturation is well documented and is crucial for stage specific regulation of spermatogenesis, removal of unviable or damaged spermatozoa, spermatozoa motility, and spermatozoa capacitation [81,82]. On the other hand, protease inhibitors attenuate unwanted proteolysis and maintain spermatozoa quiescence [81]. In our study, one of the most notable protease families present was the cathepsin proteases [83]. Cathepsins have been characterized in seminal plasma in a variety of species where they are proposed to function in membrane re-modeling and immunomodulation [83–85]. Three family members were identified in all biological replicates, including cathepsins O, K, and M. Cathepsin inhibitors cystatin and kininogen-1 were also identified **[Table S3.1]**, indicating that regulated proteolysis is likely a prominent feature in Chinook salmon seminal plasma. In addition to cathepsin proteases, proteins involved in the ubiquitin-proteasome system (UPS) were also abundant in the seminal plasma of both tactics **[Table S3.1]**. The extracellular UPS has garnered considerable attention recently in reproductive biology and has been implicated in targeting proteins for degradation for the purpose of membrane re-modeling, removal of defective

spermatozoa, and regulation of spermatozoa maturation [86]. Furthermore, there is growing evidence that measuring ubiquitin levels in seminal plasma is a viable clinical biomarker assay for infertility in humans, where higher levels correlate inversely with spermatozoa numbers, motility, and proper morphology [87]. Although we detected ubiquitin levels to be similar in our study **[Table S3.1]**, there were UPS associated enzymes whose levels were different between the tactics. Most notably, the E3 ligase KEG-like enzyme was over two-fold higher in hooknose males **[Figure 3.7A, and Table 3.1]**. As E3 ligase enzymes are responsible for recognizing targets and presenting them to the ubiquitin conjugating machinery, elevated levels suggest that hooknose males may display a greater propensity to perform UPS-mediated activities. Supporting this hypothesis was the observation that the ubiquitin carboxy-terminal hydrolase 44-like enzyme, which is predicted to function as a deubiquitinase, was 1.8 fold higher in jack males than hooknose males. As the UPS system can exert multiple effects on spermatozoa, it will be interesting to examine how these protein differences in UPS-related enzymes affect sperm competition in the alternative reproductive tactics of Chinook salmon.

In conclusion, utilizing a high throughput label-free quantitative proteomics approach that incorporated ion mobility separation and data-independent acquisition, we identified known and novel seminal plasma proteins from Chinook salmon alternative reproductive tactics. Moreover, for the first time, quantitative proteomic differences between alternative reproductive tactics (hooknose and jack males) in seminal plasma were elucidated. Many of the protein abundance differences offer insight into the divergent reproductive strategies displayed by the Chinook salmon male tactics. Future work will now be centered on examining the physiological significance of the differential protein components and their mechanisms of action in relation to competitive and non-competitive fertilization success.

3.4 – References

[1] R. Billard, Spermatogenesis and spermatology of some teleost fish species, Reprod Nutr Dev 26 (1986) 877–920.

[2] P. Stockley, M.J. Gage, G.A. Parker, A.P. Moller, Sperm competition in fishes: the evolution of testis size and ejaculate characteristics, Am Nat 149 (1997) 933–954.

[3] M. Taborsky, R.F. Oliveira, H.J. Brockmann, The evolution of alternative reproductive tactics: concepts and questions, in: R.F. Oliveira, M. Taborsky, J.H. Brockmann (Eds.), Alternative Reproductive Tactics: An Integrative Approach, Cambridge University Press, Cambridge (2008) 1–21.

[4] B.D. Neff, F. Peng, M.R. Gross, Sperm investment and alternative mating tactics in bluegill sunfish (Lepomis macrochirus), Behav Ecol 14 (2003) 634–641.

[5] G.A. Parker, Sperm competition games - sneaks and extra-pair copulations, P Roy Soc Biol Sci 242 (1990) 127–133.

[6] M.J.G. Gage, P. Stockley, G.A. Parker, Effects of alternative male mating strategies on characteristics of sperm production in the Atlantic salmon (Salmo salar): theoretical and empirical investigations, Philos T Roy Soc B 350 (1995) 391–399.

[7] T.V. Vladic, T. Jarvi, Sperm quality in the alternative reproductive tactics of Atlantic salmon: the importance of the loaded raffle mechanism, Proc Biol Sci 268 (2001) 2375–2381.

[8] B. Berejikian, D.M. Van Doornik, J.J. Atkins, Alternative male reproductive phenotypes affect offspring growth rates in Chinook Salmon, T Am Fish Soc 140 (2011) 1206–1212.

[9] D.D. Heath, R.H. Devlin, J.W. Heath, G.K. Iwama, Genetic, environmental and interaction effects on the incidence of jacking in Oncorhynchus-tshawytscha (Chinook Salmon), Heredity 72 (1994) 146–154.

[10] E.W. Flannery, I.A.E. Butts, M. Slowinska, A. Ciereszko, T.E. Pitcher, Reproductive investment patterns, sperm characteristics, and seminal plasma physiology in alternative reproductive tactics of Chinook salmon (Oncorhynchus tshawytscha), Biol J Linn Soc 108 (2013) 99–108.

[11] I.A. Butts, O.P. Love, M. Farwell, T.E. Pitcher, Primary and secondary sexual characters in alternative reproductive tactics of Chinook salmon: associations with androgens and the maturation-inducing steroid, Gen Comp Endocrinol 175 (2012) 449–456.

[12] R.R. Snook, Sperm in competition: not playing by the numbers, Trends Ecol Evol 20 (2005) 46–53.

[13] S.A. Ramm, L. McDonald, J.L. Hurst, R.J. Beynon, P. Stockley, Comparative proteomics reveals evidence for evolutionary diversification of rodent seminal fluid and its functional significance in sperm competition, Mol Biol Evol 26 (2009) 189–198.

[14] S.A. Ramm, D.A. Edward, A.J. Claydon, D.E. Hammond, P. Brownridge, J.L. Hurst, R.J. Beynon, P. Stockley, Sperm competition risk drives plasticity in seminal fluid composition, BMC Biol 13 (2015) 1–18.

[15] L. Locatello, F. Poli, M.B. Rasotto, Tactic-specific differences in seminal fluid influence sperm performance, Proc R Soc Lond B Biol Sci 280 (2013) 1–6.

[16] J.C. Silva, M.V. Gorenstein, G.Z. Li, J.P. Vissers, S.J. Geromanos, Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition, Mol Cell Proteomics 5 (2006) 144–156.

[17] R.P. Dator, K.W. Gaston, P.A. Limbach, Multiple enzymatic digestions and ion mobility separation improve quantification of bacterial ribosomal proteins by data independent acquisition liquid chromatography–mass spectrometry, Anal Chem 86 (2014) 4264–4270.

[18] N.J. Bond, P.V. Shliaha, K.S. Lilley, L. Gatto, Improving qualitative and quantitative performance for MS(E)-based label-free proteomics, J Proteome Res 12 (2013) 2340–2353.

[19] H. Niksirat, A. Kouba, P. Kozák, Post-mating morphological changes in the spermatozoon and spermatophore wall of the crayfish Astacus leptodactylus: insight into a non-motile spermatozoon, Anim Reprod Sci 149 (2014) 325–334.

[20] H. Niksirat, P. James, L. Andersson, A. Kouba, P. Kozák, Label-free protein quantification in freshly ejaculated versus post-mating spermatophores of the noble crayfish Astacus astacus, J Proteome 123 (2015) 70–77.

[21] S. Ghisaura, B. Loi, G. Biosa, M. Baroli, D. Pagnozzi, T. Roggio, S. Uzzau, R. Anedda, M.F. Addis, Proteomic changes occurring along gonad maturation in the edible sea urchin Paracentrotus lividus, J Proteome 144 (2016) 63–72.

[22] M.A. Dietrich, G.J. Dietrich, A. Mostek, A. Ciereszko, Motility of carp spermatozoa is associated with profound changes in the sperm proteome, J Proteome 138 (2016) 124–135.

[23] S.C. Stephen, Salmonine Introductions to the Laurentian Great Lakes, Canadian

Special Publication of Fisheries and Aquatic Sciences, No. 132, NRC Research Press (2001) 36–47.

[24] T.E. Pitcher, B.D. Neff, MHC class IIB alleles contribute to both additive and nonadditive genetic effects on survival in Chinook salmon, Mol Ecol 15 (2006) 2357–2365.

[25] T.E. Pitcher, B.D. Neff, Genetic quality and offspring performance in Chinook salmon: implications for supportive breeding, Conserv. Genet 8 (2007) 607–616.

[26] F. Finamore, F. Priego-Capote, F. Gluck, A. Zufferey, P. Fontana, J.-C. Sanchez, Impact of high glucose concentration on aspirin-induced acetylation of human serum albumin: an in vitro study, EuPA Open Proteomics 3 (2014) 100–113.

[27] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J R Stat Soc Ser B Methodol. 57 (1995) 289–300.

[28] R Core Team, R: A Language and Environment for Statistical Computing, https:// www.r-project.org 2016 (accessed 16/07/07).

[29] H. Wickham, ggplot: Elegant Graphics for Data Analysis, Springer-Verlag, New York, 2009.

[30] P.V. Shliaha, N.J. Bond, L. Gatto, K.S. Lilley, Effects of traveling wave ion mobility separation on data independent acquisition in proteomics studies, J Proteome Res 12 (2013) 2323–2339.

[31] K.A. Naish, R.B. Phillips, M.S. Brieuc, L.R. Newton, A.E. Elz, L.K. Park, Comparative genome mapping between Chinook salmon (Oncorhynchus tshawytscha) and rainbow trout (O. mykiss) based on homologous microsatellite loci, G3: Genes|Genomes|Genetics 3 (2013) 2281–2288.

[32] P.J. Lester, S.J. Hubbard, Comparative bioinformatic analysis of complete proteomes and protein parameters for cross-species identification in proteomics, Proteomics 2 (2002) 1392–1405.

[33] A.J. Liska, A. Shevchenko, Expanding the organismal scope of proteomics: cross-species protein identification by mass spectrometry and its implications, Proteomics 3 (2003) 19–28.

[34] B. Habermann, J. Oegema, S. Sunyaev, A. Shevchenko, The power and the limitations of cross-species protein identification by mass spectrometry-driven sequence similarity searches, Mol Cell Proteomics 3 (2004) 238–249.

[35] J.G. Burniston, J. Connolly, H. Kainulainen, S.L. Britton, L.G. Koch, Label-free profiling of skeletal muscle using high-definition mass spectrometry, Proteomics 14 (2014) 2339–2344.

[36] B. Pilch, M. Mann, Large-scale and high-confidence proteomic analysis of human seminal plasma, Genome Biol 7 (2006) 1–10.

[37] J. Nynca, G.J. Arnold, T. Frohlich, K. Otte, F. Flenkenthaler, A. Ciereszko, Proteomic identification of rainbow trout seminal plasma proteins, Proteomics 14 (2014) 133–140.

[38] M.A. Dietrich, G.J. Arnold, J. Nynca, T. Frohlich, K. Otte, A. Ciereszko, Characterization of carp seminal plasma proteome in relation to blood plasma, J. Proteome 98 (2014) 218–232.

[39] C.L. Harris, M. Mizuno, B.P. Morgan, Complement and complement regulators in the male reproductive system, Mol Immunol. 43 (1–2) (2006) 57–67.

[40] A. Poiani, Complexity of seminal fluid: a review, Behav Ecol Sociobiol 60 (2006) 289–310.

[41] M.A. Dietrich, M. Adamek, B. Bilinska, A. Hejmej, D. Steinhagen, A. Ciereszko, Characterization, expression and antibacterial properties of apolipoproteins A from carp (Cyprinus carpio L.) seminal plasma, Fish Shellfish Immunol. 41 (2014) 389–401.

[42] F.H. Pierik, J.T. Vreeburg, T. Stijnen, F.H. De Jong, R.F. Weber, Serum inhibin B as a marker of spermatogenesis, J Clin Endocrinol Metab 83 (1998) 3110–3114.

[43] A.E. O'Connor, D.M. De Kretser, Inhibins in normal male physiology, Semin Reprod Med 22 (2004) 177–185.

[44] S.I. Buschow, B.W. van Balkom, M. Aalberts, A.J. Heck, M. Wauben, W. Stoorvogel, MHC class II-associated proteins in B-cell exosomes and potential functional implications for exosome biogenesis, Immunol Cell Biol 88 (2010) 851–856.

[45] N.L. Anderson, M. Polanski, R. Pieper, T. Gatlin, R.S. Tirumalai, T.P. Conrads, T.D. Veenstra, J.N. Adkins, J.G. Pounds, R. Fagan, A. Lobley, The human plasma proteome: a nonredundant list developed by combination of four separate sources, Mol Cell Proteomics 3 (2004) 311–326.

[46] H.J. Glander, J. Kratzsch, C. Weisbrich, G. Birkenmeier, Insulin-like growth factor-I and alpha(2)-macroglobulin in seminal plasma correlate with semen quality, Hum Reprod 11 (1996) 2454–2460.

[47] A.A. Rehman, H. Ahsan, F.H. Khan, Alpha-2-Macroglobulin: a physiological guardian, J Cell Physiol 228 (2013) 1665–1675.

[48] J.M. Vaughan, W.W. Vale, Alpha 2-macroglobulin is a binding protein of inhibin and activin, Endocrinology 132 (1993) 2038–2050.

[49] G. Raposo, W. Stoorvogel, Extracellular vesicles: exosomes, microvesicles, and friends, J Cell Biol 200 (2013) 373–383.

[50] J.M. Street, P.E. Barran, C.L. Mackay, S. Weidt, C. Balmforth, T.S. Walsh, R.T. Chalmers, D.J. Webb, J.W. Dear, Identification and proteomic profiling of exosomes in human cerebrospinal fluid, J Transl Med 10 (2012) 5.

[51] D. Duijvesz, K.E. Burnum-Johnson, M.A. Gritsenko, A.M. Hoogland, M.S. Vredenbregt-van den Berg, R. Willemsen, T. Luider, L. Pasa-Tolic, G. Jenster, Proteomic profiling of exosomes leads to the identification of novel biomarkers for prostate cancer, PLoS One 8 (2013), e82589.

[52] P.A. Gonzales, T. Pisitkun, J.D. Hoffert, D. Tchapyjnikov, R.A. Star, R. Kleta, N.S. Wang, M.A. Knepper, Large-scale proteomics and phosphoproteomics of urinary exosomes, J Am Soc Nephrol 20 (2009) 363–379.

[53] A. Bobrie, M. Colombo, G. Raposo, C. Thery, Exosome secretion: molecular mechanisms and roles in immune responses, Traffic 12 (2011) 1659–1668.

[54] X. Zhao, Y. Wu, J. Duan, Y. Ma, Z. Shen, L. Wei, X. Cui, J. Zhang, Y. Xie, J. Liu, Quantitative proteomic analysis of exosome protein content changes induced by hepatitis B virus in Huh-7 cells using SILAC labeling and LC-MS/MS, J Proteome Res. 13 (2014) 5391–5402.

[55] R. Sullivan, F. Saez, J. Girouard, G. Frenette, Role of exosomes in sperm maturation during the transit along the male reproductive tract, Blood Cells Mol Dis 35 (2005) 1–10.

[56] K.G. Ronquist, B. Ek, A. Stavreus-Evers, A. Larsson, G. Ronquist, Human prostasomes express glycolytic enzymes with capacity for ATP production, Am J Physiol Endocrinol Metab 304 (2013) E576–E582.

[57] L.L. Piehl, M.L. Fischman, U. Hellman, H. Cisale, P.V. Miranda, Boar seminal plasma exosomes: effect on sperm function and protein identification by sequencing, Theriogenology 79 (2013) 1071–1082.

[58] M. Alenquer, M.J. Amorim, Exosome biogenesis, regulation, and function in viral infection, Viruses 7 (2015) 5066–5083.

[59] I.A. Rooney, J.P. Atkinson, E.S. Krul, G. Schonfeld, K. Polakoski, J.E. Saffitz, B.P.

Morgan, Physiologic relevance of the membrane attack complex inhibitory protein CD59 in human seminal plasma: CD59 is present on extracellular organelles (prostasomes), binds cell membranes, and inhibits complement-mediated lysis, J Exp Med 177 (1993) 1409–1420.

[60] A.S. Block, S. Saraswati, C.F. Lichti, M. Mahadevan, A.B. Diekman, Co-purification of Mac-2 binding protein with galectin-3 and association with prostasomes in human semen, Prostate 71 (2011) 711–721.

[61] A. Clayton, C.L. Harris, J. Court, M.D. Mason, B.P. Morgan, Antigen-presenting cell exosomes are protected from complement-mediated lysis by expression of CD55 and CD59, Eur J Immunol 33 (2003) 522–531.

[62] O. Serang, J. Paulo, H. Steen, J.A. Steen, A non-parametric cutout index for robust evaluation of identified proteins, Mol Cell Proteomics 12 (2013) 807–812.

[63] L. Huang, S.I. Wickramasekara, T. Akinyeke, B.S. Stewart, Y. Jiang, J. Raber, C.S. Maier, Ion mobility-enhanced MSE-based label-free analysis reveals effects of low-dose radiation post contextual fear conditioning training on the mouse hippocampal proteome, J Proteome 140 (2016) 24–36.

[64] C. Selby, Sex hormone binding globulin: origin, function and clinical significance, Ann Clin Biochem 27 (1990) 532–541.

[65] L. Lazaros, N. Xita, A. Kaponis, K. Zikopoulos, N. Sofikitis, I. Georgiou, Evidence for association of sex hormone-binding globulin and androgen receptor genes with semen quality, Andrologia 40 (2008) 186–191.

[66] C.A. Heinlein, C. Chang, The roles of androgen receptors and androgen-binding proteins in nongenomic androgen actions, Mol Endocrinol 16 (2002) 2181–2187.

[67] S. Tardif, O.A. Madamidola, S.G. Brown, L. Frame, L. Lefievre, P.G. Wyatt, C.L. Barratt, S.J. Martins Da Silva, Clinically relevant enhancement of human sperm motility using compounds with reported phosphodiesterase inhibitor activity, Hum Reprod 29 (2014) 2123–2135.

[68] L. Lefievre, E. De Lamirande, C. Gagnon, The cyclic GMP-specific phosphodiesterase inhibitor, sildenafil, stimulates human sperm motility and capacitation but not acrosome reaction, J Androl 21 (2000) 929–937.

[69] D.D. Hoskins, M.L. Hall, D. Munsterman, Induction of motility in immature bovine spermatozoa by cyclic AMP phosphodiesterase inhibitors and seminal plasma, Biol Reprod 13 (1975) 168–176.

[70] M. Morisawa, K. Ishida, Short-term changes in levels of cyclic AMP, adenylate

cyclase, and phosphodiesterase during the initiation of sperm motility in rainbow trout, J Exp Zool 242 (1987) 199–204.

[71] P. Dzeja, A. Terzic, Adenylate kinase and AMP signaling networks: metabolic monitoring, signal communication and body energy sensing, Int J Mol Sci 10 (2009) 1729–1772.

[72] F. Lahnsteiner, B. Berger, T. Weismann, R.A. Patzner, Motility of spermatozoa of Alburnus alburnus (Cyprinidae) and its relationship to seminal plasma composition and sperm metabolism, Fish Physiol Biochem 15 (1996) 167–179.

[73] F. Lahnsteiner, B. Berger, T. Weismann, R.A. Patzner, Determination of semen quality of the rainbow trout, Oncorhynchus mykiss, by sperm motility, seminal plasma parameters, and spermatozoal metabolism, Aquaculture 163 (1998) 163–181.

[74] W. Cao, L. Haig-Ladewig, G.L. Gerton, S.B. Moss, Adenylate kinases 1 and 2 are part of the accessory structures in the mouse sperm flagellum, Biol Reprod 75 (2006) 492–500.

[75] C. Burgos, C. Maldonado, N.M. Gerez de Burgos, A. Aoki, A. Blanco, Intracellular localization of the testicular and sperm-specific lactate dehydrogenase isozyme C4 in mice, Biol Reprod 53 (1995) 84–92.

[76] F. Odet, S.A. Gabel, J. Williams, R.E. London, E. Goldberg, E.M. Eddy, Lactate dehydrogenase C and energy metabolism in mouse sperm, Biol Reprod 85 (2011) 556–564.

[77] J.L. Fitzpatrick, P.M. Craig, C. Bucking, S. Balshine, C.M. Wood, G.B. McClelland, Sperm performance under hypoxic conditions in the intertidal fish Porichthys notatus, Can J Zool 87 (2009) 464–469.

[78] W.C. Ford, Regulation of sperm function by reactive oxygen species, Hum Reprod Update 10 (2004) 387–399.

[79] L. Yan, J. Liu, S. Wu, S. Zhang, G. Ji, A. Gu, Seminal superoxide dismutase activity and its relationship with semen quality and SOD gene polymorphism, J Assist Reprod Genet. 31 (2014) 549–554.

[80] T. Kobayashi, T. Miyazaki, M. Natori, S. Nozawa, Protective role of superoxide dismutase in human sperm motifity: superoxide dismutase activity and lipid peroxide in human seminal plasma and spermatozoa, Hum Reprod 6 (1991) 987–991.

[81] B.A. Laflamme, M.F. Wolfner, Identification and function of proteolysis regulators in seminal fluid, Mol Reprod Dev 80 (2013) 80–101.

[82] S.S. Chung, L.J. Zhu, M.Y. Mo, B. Silvestrini, W.M. Lee, C.Y. Cheng, Evidence for crosstalk between Sertoli and germ cells using selected cathepsins as markers, J Androl 19 (1998) 686–703.

[83] V. Turk, V. Stoka, O. Vasiljeva, M. Renko, T. Sun, B. Turk, D. Turk, Cysteine cathepsins: from structure, function and regulation to new frontiers, Biochim Biophys Acta 1824 (2012) 68–88.

[84] S. Conus, H.U. Simon, Cathepsins and their involvement in immune responses, Swiss Med Wkly 140 (2010) w13042.

[85] C. Bonnans, J. Chou, Z. Werb, Remodelling the extracellular matrix in development and disease, Nat Rev Mol Cell Biol 15 (2014) 786–801.

[86] R. Bose, G. Manku, M. Culty, S.S. Wing, Ubiquitin-proteasome system in spermatogenesis, Adv Exp Med Biol 759 (2014) 181–213.

[87] P. Sutovsky, R. Hauser, M. Sutovsky, Increased levels of sperm ubiquitin correlate with semen quality in men from an andrology laboratory clinic population, Hum Reprod 19 (2004) 628–638.

CHAPTER 4- PATHOLOGIC ALTERATIONS IN THE PROTEOME OF SYNAPTOSOMES FROM A MOUSE MODEL OF SPINAL MUSCULAR ATROPHY

4.1 - Introduction

Spinal muscular atrophy (SMA) is the leading hereditary cause of death in infants[1]. It is characterized by progressive weakness and atrophy of proximal muscles [1]. SMA is an autosomal recessive hereditary disease and is caused by mutations in the *survival motor neuron 1* (*SMN1*) gene [2]. Type I SMA is the most severe form of the disease and accounts for about 50% of all patients with SMA [3,4]. In this type, the disease starts early in life and usually results in death before the age of two. In postmortem examination, significant loss of lower motor neurons in the ventral horn of the spinal cord and atrophy of skeletal muscles are observed [5]. The role of SMN, the protein product of *SMN1*, in the assembly of protein-RNA complexes (snRNPs) and pre-mRNA splicing is well established [6]. The depletion of SMN to very low levels is not lethal but results in the SMA phenotype in humans and in other animals [7,8].

Neuromuscular junctions (NMJs) are among the most studied structures in mouse models of SMA. In general, NMJs develop normally during embryonic life in these mice, but undergo degenerative changes along the course of the disease. The degree of the NMJ pathology increases with progress of SMA phenotype in these mice [9-11]. These degenerative changes include accumulation of neurofilaments within pre-synaptic areas, reduced size and delayed maturation of motor endplates, and eventual denervation of motor endplates. Some of the pathologies involving NMJs (e.g. neurofilament accumulation) occur before any overt symptoms in affected mice [10]. In addition, Dachs et al. showed

that the expression of some of the presynaptic proteins (including Rab3 and calcitonin gene-related peptide) is reduced in NMJs of a severe SMA mouse model before the onset of the disease [12]. Other studies also showed several ultrastructural impairments of NMJs including decreased density of synaptic vesicles in the active zone, reduced number of presynaptic mitochondria, and delayed switch of postsynaptic acetylcholine receptor subunits [13]. In accordance with these findings, electrophysiological studies on NMJs of SMA mice also showed a variety of abnormalities including reduction of end-plate potentials, decrease in quantum content and reduction in release probability [14]. Several NMJ abnormalities including aberrant synaptic vesicles and accumulation of neurofilaments in the motor nerve terminals have been reported in samples of human SMA embryos [15]. The authors concluded that the pathologic changes in severe forms of SMA start during embryonic development and a defect in maintaining the initial innervations of motor endplates is an early event in SMA pathogenesis [15].

Other studies using fly and mouse models have suggested that the impairment of neural circuits within the spinal cord might contribute to the motor deficits of SMA [16-18]. Using a severe mouse model of SMA, Mentis et al. studied the monosynaptic connections between motor neurons and muscle spindles [17]. They showed that even before the onset of the disease, the response of motor neurons to stimulation of dorsal root (afferent fibers) is significantly decreased in SMA mice [17]. They further examined lumbar sections of spinal cords in these animals and found that the number of proprioceptive synapses (marked by an antibody against VGlut1) on the soma and dendrites of lower motor neurons is remarkably reduced. Moreover, Wishart et al. observed protein differences using isobaric labeling and quantitative proteomics on synaptosome

preparations from the hippocampus from severe SMA mice at postnatal day 1 (P1) [19]. Thus, the emerging view is that synapses and axons play an important role in the pathogenesis of a number of neurodegenerative disorders. As a result, treatment strategies targeting neuroprotection, and in particular synapses and axons become a viable option. Here, we have used the *Smn*$^{2B/-}$ mouse model of severe SMA to explore aberrant molecular events at central synapses accompanying disease pathogenesis and to identify proteins and pathways impacting neurodegeneration.

4.2 - Experimental Section

Mouse Maintenance and Handling

C57BL/6 *Smn*$^{+/-}$ and control mice were purchased from the Jackson Laboratory as follows: C57BL/6J (#000664), and B6.129P2(Cg)-Smn1<tm1Msd>/J (#010921) (Bar Harbor, Maine, USA). *Smn*$^{2B/2B}$ mice were previously generated in our laboratory and have been maintained on a C57BL/6 background [20,21]. All animals were handled according to institutional guidelines (Animal Care and Veterinary Services, University of Ottawa). The humane endpoints included severe dehydration, hypothermia or dragging of the hind limbs.

Preparation of Synaptosome Fractions

Based on methods described by Dunkley et al. (2008), non-continuous Percoll gradients were used to obtain highly pure and enriched synaptosome fractions from mouse spinal cords and brain cortices [22]. Briefly, mice were euthanized, and their cortices and spinal cords were dissected immediately. Tissue samples were rinsed in ddH2O and were immediately homogenized in 6 mL of a sucrose buffer (containing 320 mM sucrose, 5 mM Tris-HCl pH 7.5 and 0.1 mM EDTA) using a 7 mL Dounce tissue grinder (10 strokes loose and 10 strokes tight pestles). The homogenates were centrifuged at 1,100 g (10 min at 4°C) and the supernatants were transferred to new tubes and were centrifuged at 20,000 g (10 min at 4°C). The supernatants were discarded, and the pellets were resuspended in 2 mL of gradient medium (GM) buffer (containing 250 mM sucrose, 5 mM Tris-HCl pH 7.5 and 0.1 mM EDTA) and loaded gently on top of non-continuous Percoll gradients. The gradient tubes were centrifuged at 45,000 g (5 min at maximum speed, 4°C) using a fixed angle 70.1 Ti rotor (Beckman). The 3rd and 4th fraction of each tube was aspirated carefully using

a Pasteur pipette and transferred to new tubes and diluted 10 times by GM buffer and centrifuged at 45,000 x g (5 min at maximum speed, 4°C). The supernatants were discarded, and the pellets were resuspended in 2 mL of GM buffer and transferred to microcentrifuge tubes and centrifuged in a benchtop machine at maximum speed (5 min at maximum speed, 4°C). The retrieved pellets were considered as synaptosome fractions and were used for further analysis.

Electron Microscopy

Transmission electron microscopy was used to determine the integrity and quality of prepared synaptosomes. Fresh synaptosome fractions prepared from mouse cortices and spinal cords were incubated with Karnovsky's fixative buffer (4% paraformaldehyde, 2% glutaraldehyde and 0.1 M sodium cacodylate in PBS, pH 7.4) for at least 1 h at room temperature. Samples were centrifuged at 20,000 g for 1 min and the pellets were washed three times in 0.1 M sodium cacodylate buffer for 10 min. Samples were post-fixed with 1% osmium tetroxide in 0.1 M sodium cacodylate buffer for 2 h and washed three times in distilled water for 5 min. After the last wash, the pellets were dehydrated in increasing concentrations of ethanol (50%, 75% and 90% 15 min each), washed for 10 min in a ethanol/acetone (50:50) solution followed by centrifugation at 20,000 g for 1 min, and then were washed for 15 min in 100% acetone followed by centrifugation at 20,000 g for 1 min. The samples were infiltrated first in 30% Spurr resin/acetone overnight, then in 50% Spurr resin/acetone for 6 h, and finally in fresh 100% Spurr resin overnight (all infiltration steps were performed on a rotator at low speed). Synaptosomes were embedded in fresh liquid Spurr resin and then polymerized overnight at 70ºC. The specimens were sectioned using an ultramicrotome at 80 nm thickness. The ultrathin sections were collected onto 200-mesh

copper grids and stained using 2% aqueous uranyl acetate and then Reynold's lead citrate. Several electron micrographs were prepared from the stained sections using a transmission electron microscope (Hitachi 7100) at 10,000x and 100,000x magnification, and then ultrastructural analysis was performed by visual examination.

Immunoblot Analysis

Synaptosome pellets were incubated with RIPA buffer (Cell Signaling Technology) for 10 min on ice. Protein concentrations of homogenates were measured using a Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific) and 10 µg total protein from each sample was incubated with Laemmli loading buffer at 95°C for 5 min. Samples were separated on 10 or 12% SDS-PAGE gels and then electrotransferred onto Immobilon-FL membranes (EMD Millipore). Before blocking, the blotted membranes were stained for total protein using Sypro Ruby staining reagent (Life Technologies) and scanned by a Chemidoc-IT imager (UVP). The following primary antibodies were used to probe the membranes: mouse anti-histone H3 1:2000 (Thermo), mouse anti-β-actin (1:10000), mouse anti MBP and anti GFAP (1:2000) (Millipore), rabbit anti-SV2 (1:2000) (DSHB), rabbit anti- Synaptophysin, PSD-95, COX4 and SDHA (1:2000) (Cell Signaling Technologies). Membranes then were incubated with IRDye fluorescent conjugated (LiCOR) (1:5000) and were developed using an Odyssey CLx scanning machine. The images were quantified using Image Studio 4.0 (LiCOR). Total protein of each lane was used to normalize the signals within that lane.

In-Gel Trypsin Digestion and Mass Spectrometry Sample Preparation

Synaptosome fractions from six biological replicates; three control samples and three *Smn*$^{2B/-}$ samples, were resuspended in 6X SDS loading dye (10% SDS w/v, 0.5 M Tris-HCl pH 6.8, 40% glycerol (v/v), and 30 mM DTT) and boiled. Samples were separated on 15% SDS-PAGE gels and visualized using Imperial Coomassie stain (Thermo Fisher Scientific) following the manufacturer's protocol. Individual lanes were excised and further dissected into eight parts. Each gel piece was placed into individual siliconized tubes (Fisher Scientific) and destained as previously described prior to in-gel trypsin digestion [23]. Mass spectrometry grade trypsin (Promega) was prepared in 50 mM NH_4HCO_3 (Thermo Fisher Scientific) and made up to a final concentration of 0.04 µg/µL. Approximately 0.8 µg of trypsin was added to each gel piece, which was incubated at 37°C overnight. Trypsin digestion was quenched with 5% formic acid (Sigma). Peptides were extracted using 60% acetonitrile (Burdick and Jackson), 1% formic acid in HPLC grade water (Burdick and Jackson). Peptide extracts for each biological replicate were pooled together to yield 4 fractions and concentrated to dryness using vacuum centrifugation. Samples were reconstituted in 0.1% formic acid solution and Hi3 standard peptides (Waters) (12.5 fmol/µL) were added for relative protein quantitation.

Mass Spectrometry Analysis

Generated in-gel tryptic peptides from each of the four fractions per biological replicate were injected on a 1.8 µm HSS T3 75 µm x 150 mm reverse phase column (Waters) using a nanoAcquity UPLC system (Waters) at a flow rate of 0.3 µL/min. Mobile phase A (water with 0.1% formic acid) and mobile phase B (acetonitrile with 0.1% formic

acid) were used to equilibrate the column and load the peptides onto the column at a ratio of 97:3 (A:B). Elution of the peptides was accomplished using a 85 min acetonitrile gradient (3-30% B for 60 min, 30-50% B for 15 min, 85% B for 10 min). Peptides were electrosprayed directly into a SYNAPT G2-Si mass spectrometer (Waters) using a cone voltage of 30 V and a capillary voltage of 3 kV. The instrument was operated in data independent acquisition mode with the ion mobility chamber activated. Low energy scans were acquired at 4 eV while high energy scans were acquired using a 20-45 eV ramp in positive resolution mode scanning from 50-2000 m/z with a scan rate of 0.8 s. Data was externally calibrated using [Glu1]- fibrinopeptide B (50 fmol/µL) in the lockmass channel. Data was collected using MassLynx (version 4.1) and Progenesis QI Nonlinear Dynamics) was used for chromatography alignment, normalization, and peptide identification analysis using the *Mus musculus* subset of the UniprotKB/SwissProt database (17016 proteins, downloaded January 2018). The raw data was processed using the following parameters: a low energy noise reduction threshold of 135, a high-energy threshold of 30, and an intensity threshold of 750. The data was lock mass corrected post acquisition. Protein identification parameters included a minimum of three fragment ions per peptide, a minimum of seven fragment ion matches per protein, a maximum of two missed cleavages for trypsin, and a false discovery rate below 1% was achieved using a decoy database. Variable modification for oxidation of methionine was also specified. Three technical replicates were acquired for each of the four fractions per biological replicate (72 measurements total) and grouped into the comparative categories (control vs $Smn^{2B/-}$) in Progenesis QI. Quantitation was determined using the three most abundant peptides per protein following conflict resolution and elimination of protein contaminants. The mass spectrometry proteomics data have been

deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD012850.

Label-Free Quantitative Data Analysis

Quantitative data were analysed using "R" statistical software (Version 3.5.1). Statistical significance, relating to the differences in protein abundance (fmols) between control and SMA mice, were evaluated using a student's t-test and was accepted if P-values were <0.05. Prior to performing the student's t-test, the average protein abundance was calculated from three technical replicates per biological replicate of control and SMA mice. The reliability of quantitation was standardized based on the detection of at least three unique peptides per protein. To ensure statistical relevance, changes in protein abundance was considered significant only if the student's t-test was <0.05, the ANOVA score was <0.05, and the fold-change was >1.3. Statistical data from all quantitated proteins was used to generate a volcano plot. To represent significant changes in protein abundance, the student's t-test scores were converted to $-\log_{10}$ P-values (y-axis) and plotted in relation to $\log_2$fold-change (x-axis) using "ggplot2" and "ggrepel" packages in "R".

Bioinformatics Data Analysis

Proteomic data were processed using *The Database for Annotation, Visualization and Integrated Discovery* (*DAVID* v6.8), available at http://david.ncifcrf.gov, to investigate the enrichment and annotation of synaptosome proteins to cellular component, biological process, molecular function, and KEGG pathways [24,25]. A Fischer Exact test was performed by *DAVID* to determine significance of protein enrichment. Enrichment of

proteins, in the categories mentioned above, was considered significant if the P-values (Fischer Exact test) were <0.05, and false-discovery rate (FDR) was <5%.

4.3 - Results and Discussion

Synaptosome Fractions Prepared from Mouse Cortices and Spinal Cords Display High Purity and Quality

To identify perturbations in molecular signaling pathways within central synapses and axons in *Smn*$^{2B/-}$ mice, we chose to assess the proteome within distal neuronal regions. To achieve this, we prepared synaptosomes, which represent synapse-enriched biochemical fractions, from P14 brain cortices and spinal cords. **Figure 4.1** shows a representative biochemical fractionation from control mice analyzed by immunoblotting. At the end of the ultracentrifugation, five distinct fractions, here labeled as F1 to F5, were collected from the Percoll gradients. Fractions F3 and F4 contained the synaptosomes as predicted from previous work [22]. These fractions showed minimum reactivity with antibodies against histone H3 (nuclear marker), myelin basic protein (MBP) (3 isoforms) (myelin marker), or glial fibrillary acidic protein (GFAP) (astrocyte marker) **[Figure 4.1]**. We also used antibodies against synaptic markers to investigate the enrichment of these proteins within the synaptosome fractions. Fractions F3 and F4 showed strong signals when antibodies against synaptic vesicle glycoprotein 2 (SV2), synaptophysin and postsynaptic density protein 95 (PSD-95) were used to probe the membranes. These results confirmed good purity and enrichment of synaptosomes within fractions F3 and F4 **[Figure 4.1]**. Fractions F3 and F4 also reacted with antibodies against 'cytochrome c oxidase 4' (COX4) and 'succinate dehydrogenase A' (SDHA), markers for mitochondria **[Figure 4.1]**. This confirms the existence of intra-synaptic mitochondria within the enriched synaptosomes as

reported by others [26,27]. However, most of the mitochondria are extra-synaptic and separate out into fraction F5 **[Figure 4.1]**. Using electron microscopy (EM), we also investigated the quality of the synaptosome fractions prepared from mouse cortices and spinal cords **[Figure 4.2]**. Fractions F3 and F4 were combined and represented the synaptosome preparation. EM showed a good integrity of synaptosome membranes and presence of synaptic vesicles and mitochondria inside the synaptosome sacs **[Figure 4.2]**, indicating the isolation of healthy and viable synaptosomes. Using immunofluorescence microscopy, we also observed similar number of synapses in lumbar spinal cords of $Smn^{2B/-}$ mice compared to control mice **[Figure S4.1]**. Together, these results confirmed that our synaptosome fractions have good quality and high enrichment of central synapses from mouse CNS tissues.

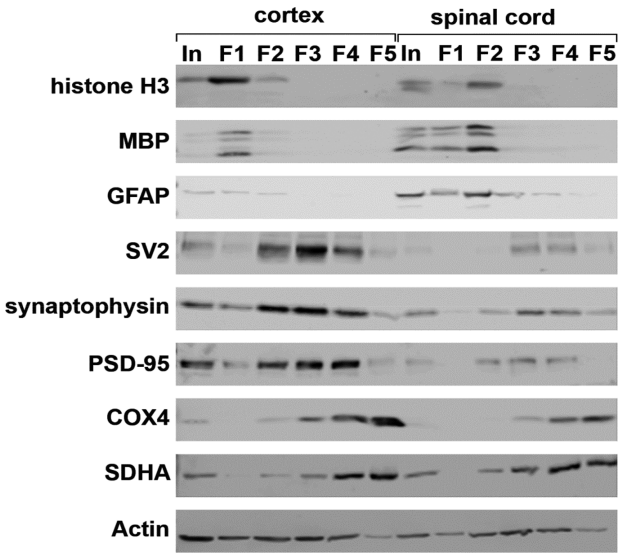


Figure 4.1 - Synaptosomal fractions prepared from control mouse cortex and spinal cord show good purity and enrichment. Using a Percoll noncontinuous gradient, homogenates of mouse cortex and spinal cord were fractionated to five distinct fractions (F1−F5, In = input). Equal protein concentrations were used from the input and each fraction to perform immunoblot analysis. Probing of the membranes with different antibodies showed a good purity and enrichment of synaptosomes within fractions F3 and F4 from mouse cortex and spinal cord. Histone H3 was used as a nuclear marker; MBP (3 isoforms) and GFAP as markers for myelin and astrocytes; actin as a marker for the cytoplasm; SV2, synaptophysin, and PSD-95 as markers for synaptic regions; and Cox4 and SDHA as markers for mitochondria. Please note that fractions F3 and F4 react with the antibodies specific for mitochondria protein. This indicates that the retrieved synaptosomes contain intrasynaptic mitochondria.

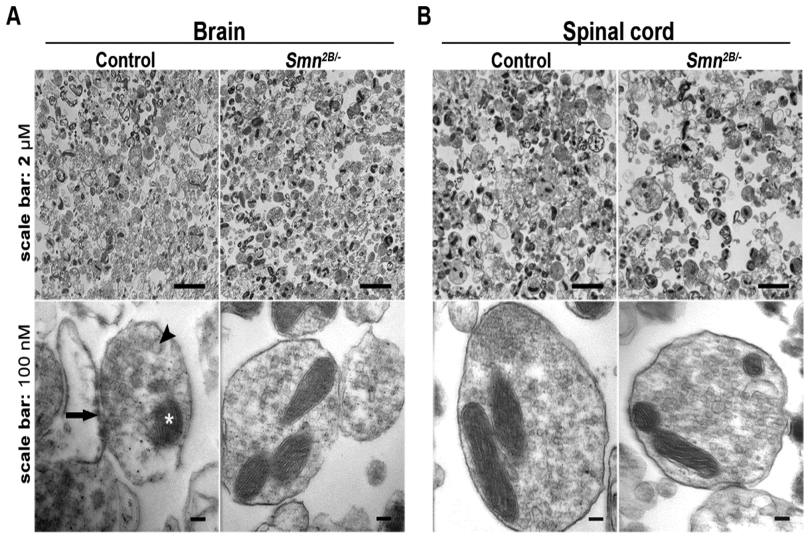


Figure 4.2 - Synaptosome fractions prepared from brain cortices and spinal cords of control and $Smn^{2B/-}$ mice show good membrane integrity and quality. Electron microscopy of synaptosome fractions prepared from mouse (A) brain cortex and (B) spinal cord confirm the integrity of the membranes of retrieved synaptosomes in control and $Smn^{2B/-}$ mice. Note the synaptosome sacs contain synaptic vesicles and intrasynaptic mitochondria (arrowhead and asterisk, respectively). Also, some synaptosomes are still attached to the postsynaptic membrane (arrow). Scale bars correspond to 2 μm (top images) and 100 nm (bottom images).

Proteome Profile of Synaptosomes Prepared from Spinal Cords of $Smn^{2B/-}$ Mice

Mass spectrometry analysis on synaptosomes isolated from mouse spinal cords was performed to investigate proteome dynamics at the motor neuron synapse in response to SMN depletion. Following synaptosome isolation, samples were prepared for mass spectrometry using an in-gel approach **[Figure 4.3]**. Identification of synaptosome proteins was accomplished using data-independent acquisition mass spectrometry (LC-MS^E). Furthermore, ion mobility separation of peptides was coupled to LC-MS^E (LC-$HDMS^E$) to improve peptide resolution [23,28]. Following data filtering to include proteins identified in all six biological replicates (3 control mice and 3 $Smn^{2B/-}$ mice), our analysis successfully identified 2030 proteins **[Table S4.1]**. Enrichment analysis, using the *DAVID* software program, was applied to evaluate synaptic-related proteins represented in our control and $Smn^{2B/-}$ samples. Indeed, our results indicate a significant enrichment of proteins associated with numerous synaptic components including synaptic vesicles, axons, dendrites, synaptic membranes, and postsynaptic density **[Table 4.1]**. The inclusion of label-free internal peptide standards (Hi3) was added to each biological replicate to aid in relative quantitation analysis (see below). Moreover, the addition of a known amount of Hi3 peptides also allows for estimation of on-column protein abundances that is useful for determining the dynamic range of our dataset [29]. Upon analyzing the abundance of synaptosome proteins on column, we observed a dynamic range spanning six orders of magnitude (0.004-140 fmols) **[Figure 4.4]**. Cytoplasmic actin, which has diverse presynaptic and postsynaptic functions, was the highest abundant protein identified in our dataset **[Table S4.1]** [30]. A mitochondrial protein involved in phospholipid metabolism, known as phosphatidylserine decarboxylase proenzyme, was identified at the low end of the spectrum [31]. Taken

together, functional evaluation of the proteomic data demonstrates the relevance of our dataset to assessing alterations in spinal cord synaptic processes in our SMA model system.

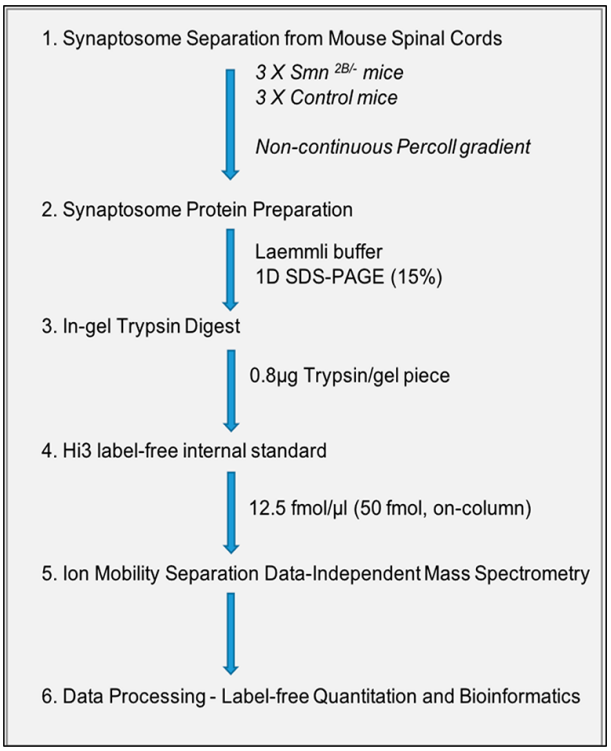


Figure 4.3 - Schematic of synaptosome experimental proteomic workflow. Spinal cord synaptosome fractions were prepared from control (n = 3) and Smn$^{2B/-}$ (n = 3) mice. Samples were subjected to SDS/PAGE analysis, in-gel trypsin digestion, and analyzed by high resolution mass spectrometry followed by statistical data processing.

Table 4.1 - DAVID Gene Ontology Enrichment Analysis of Synaptic-Related Proteins Identified in Control and Smn$^{2B/-}$ Mice.

Gene Ontology	count	*P*-value	fold enrichment	Benjamini
mitochondrion	697	3.59×10^{-291}	4.17	2.78×10^{-288}
synapse	131	3.56×10^{-26}	2.67	1.72×10^{-24}
dendrite	100	1.06×10^{-12}	2.10	1.87×10^{-11}
axon	99	5.66×10^{-21}	2.75	1.99×10^{-19}
postsynaptic density	67	2.06×10^{-15}	2.88	4.95×10^{-14}
synaptic vesicle	58	3.11×10^{-25}	4.73	1.27×10^{-23}
postsynaptic membrane	34	8.92×10^{-03}	1.58	4.32×10^{-02}
synaptic vesicle membrane	30	3.43×10^{-14}	5.06	7.18×10^{-13}
clathrin-coated pit	30	1.98×10^{-14}	5.14	4.37×10^{-13}
presynaptic membrane	27	3.53×10^{-09}	3.70	2.73×10^{-06}
synaptic membrane	13	3.44×10^{-05}	4.05	2.93×10^{-04}
neuromuscular junction	13	1.92×10^{-02}	2.09	8.46×10^{-02}

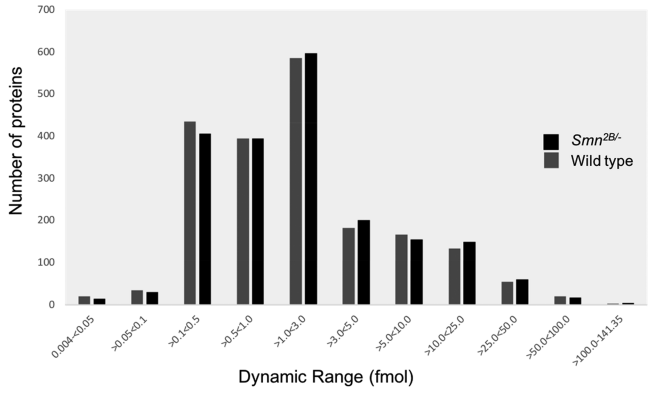


Figure 4.4 - Profile of synaptosome proteome. The dynamic range of on- column protein abundance. The distribution of proteins (y-axis) within a range of on-column protein abundances in fmol (x-axis). Gray bars represent control proteins (control), and black bars represent Smn$^{2B/-}$ mouse proteins.

Quantitative Proteomic Analysis Reveals Protein Alterations in Synaptosomes From $Smn^{2B/-}$ Mice

In attempt to disclose proteins involved in the synaptic dysfunction of motor neurons in SMA, statistical analysis of the label-free quantitative proteomic data was performed, and significant protein abundance differences in synaptosomes of control and $Smn^{2B/-}$ mice were elucidated. Our initial analysis, using a student's t-test to assess significant differences between the means of control and $Smn^{2B/-}$ mice, resulted in 170 proteins (~8% of our dataset) with P-values below 0.05 **[Table S4.1]**. Moreover, further statistical rigor assessing variance (ANOVA) and fold change (>1.3) yielded 65 proteins (~3% of our dataset) whose abundance differences were considered significantly relevant **[Table S4.1]**.

To graphically represent the statistical data, a volcano plot was constructed by inputting the $\log_{10}$ P-value as a function of $\log_2$ fold-change. Interestingly, the vast majority of the synaptosome proteins were observed at statistically similar levels between control and $Smn^{2B/-}$ mice **[Figure 4.5]**. Of the altered proteins, 58 out of the 65 proteins were observed to be significantly higher in synaptosomes from $Smn^{2B/-}$ mice, while only 7 proteins: puromycin sensitive aminopeptidase (PSA), diphosphoinositol polyphosphate phosphoryhydrolase 3-alpha (DIPP3), CD81, selenoprotein M, 60s ribosomal protein L12, alpha-actinin-3, and glucosamine-6-phosphate isomerase 2 were significantly reduced in synaptosomes of $Smn^{2B/-}$ mice **[Table S4.1]**. The clear pattern of protein alterations being elevated in SMA synaptosomes rather than reduced is suggestive of a defect in regulating the abundance of specific proteins in the synapse. Alternatively, these results may indicate that SMN depletion induces a compensatory adaptation response resulting in the observed protein alterations.

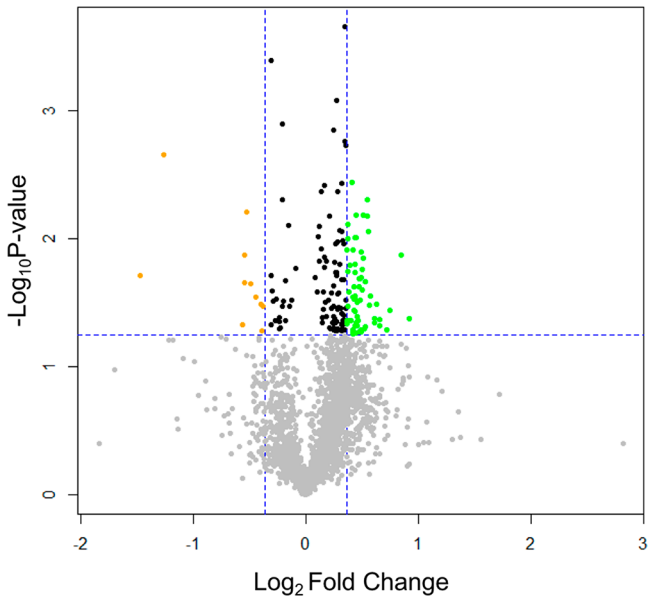


Figure 4.5 - Volcano plot depicting differentially abundant synaptosome proteins. The volcano plot displays −log10 corrected t test values (y- axis) versus log2 fold change (x-axis). Data points above horizontal dashed line (P-values below 0.05) and to the left (orange) and right (green) of the vertical dashed lines (abundance ratios above 1.3) represent proteins with statistically significant abundance differences.

To support the accuracy of the statistical analysis, we examined the synaptosome abundance levels of two actin-regulating proteins (moesin and α-actinin-3) by immunoblot analysis, whose mean differences were just below the cut-off value (P-value < 0.05) and also displayed a moderate level of variation amongst biological replicates but were concluded to be statistically significant. In addition, levels of cofilin-1 and cofilin-2 were probed to serve as negative controls as these actin-binding proteins were just above the cut-off values and whose fold-change differences were concluded to be statistically insignificant due to high variation among biological replicates **[Table S4.1]**. As shown, α-actinin-3 was confirmed to have higher protein levels in control mice **[Figure 4.6A and 4.6B]**, whereas moesin was observed at higher protein levels in the *Smn*$^{2B/-}$ mice **[Figure 4.6A and 4.6C]**. Furthermore, the protein abundance for cofilin-1and cofilin-2 did not significantly change between control and *Smn*$^{2B/-}$ mice **[Figure 4.6A and 4.6D-E]**, further confirming our mass spectrometry data analysis. The altered levels of moesin and α-actinin-3 are intriguing as impaired regulation of the actin cytoskeleton has been proposed to be a primary defect in SMA [32]. The maintenance and organization of the actin cytoskeleton is indispensable for synaptogenesis, establishing correct cell polarity, synaptic plasticity, and proper neurotransmitter release in neurons [30]. Therefore, it might be valuable to further explore the functional role of moesin and α-actinin-3 in the SMA model of motor neuron degeneration.

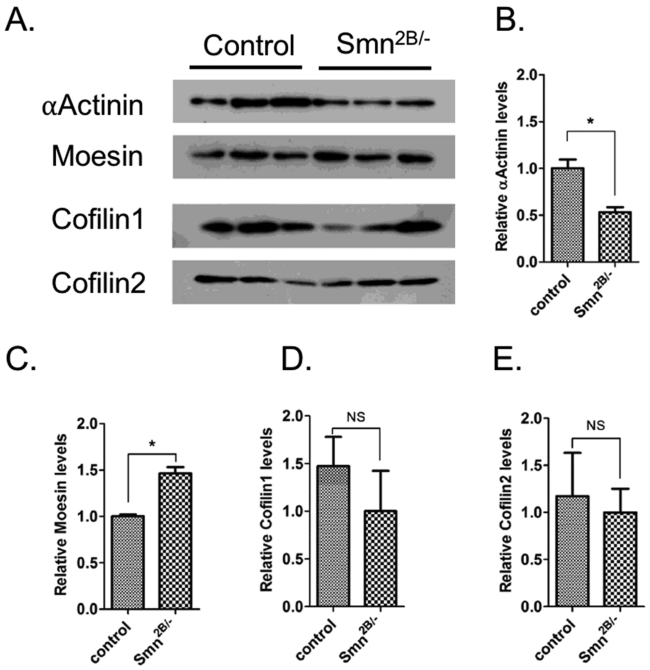


Figure 4.6 - Levels of actin cytoskeleton regulators moesin and α- actinin-3 are altered within synapses of spinal cords of Smn$^{2B/-}$ mice at a presymptomatic stage. (A) Representative images of immuno- blotting experiments on synaptosome fractions prepared from spinal cords of Smn$^{2B/-}$ mice and their control littermates at PND11. (B−E) Quantification of immunoblotting images showing a decrease in the levels of α-actinin-3, while the levels of moesin are increased in synapses of spinal cord of Smn$^{2B/-}$ mice at PND11 (n = 3; paired t test, $P < 0.05$ was considered significant). ∗ indicates statistically different, NS indicates not statistically significant. Protein levels were normalized to total protein levels as described in the **Experimental Section.**

Reduced Levels of SMN Results in Synapse Protein Imbalance Concerning Neurotransmitter Toxicity, Mitochondrial Dysfunction, and Cholesterol Metabolism

Inspection of the protein alterations identified several proteins that have previously been reported to be associated with neurological/neurodegenerative disorders **[Table 4.2]**. For example, we observed a statistically significant increase in fold-change for glycyl tRNA ligase (GlyRS) in the *Smn*$^{2B/-}$ mice **[Table 4.2]**. Mutations in the gene that encodes for GlyRS cause Charcot-Marie-Tooth type 2D (CMT2D), a progressive neuropathy that is phenotypically analogous to distal SMA type-V (SMA-DV) [33]. In CMT2D, GlyRS mutants alter its functional interactome resulting in disruption of signaling pathways critical for peripheral axon survival [34]. The abundance of transmembrane protein 106B (TMEM106B) and Annexin-A2 (AnxA2) was also higher in synaptosomes from *Smn*$^{2B/-}$ mice when compared to control mice **[Table 4.2]**. Overexpression of TMEM106B is cytotoxic, causing oxidative stress-induced cell death, and has been associated with frontotemporal lobar degeneration (FTLD), a condition where 15% of cases develop motor neuron impairments similar to amyotrophic lateral sclerosis (ALS) [35,36]. AnxA2, a Ca^{2+} binding protein known to mediate membrane organization and signal transduction, has been shown to interact and interfere with the proper localization of tau [37,38]. Because of this, AnxA2 is believed to play a neurodegenerative role in Alzheimer's disease [38]. Furthermore, AnxA2 was reported to be a direct binding partner of SMN [39]. Lastly, we observed that the protein abundance of the essential neural protein neurochondrin (NCDN) is elevated in synaptosomes from *Smn*$^{2B/-}$ mice **[Table 4.2]**. Like AnxA2, NCDN has recently been shown to associate with SMN in neurites and regulate its localization and has been proposed to be a novel therapeutic target for SMA [40].

Table 4.2 - Significant Protein Alterations in Spinal Cord Synaptosomes from Smn$^{2B/-}$ Mice.

biological function	ID	protein description	FC	*P*-value
Actin Dynamics				
higher in *Smn*$^{2B/-}$	P26041.3	moesin	1.43	0.05
lower in *Smn*$^{2B/-}$	O88990.1	alpha-actinin-3	1.31	0.05
Connection to Motor Neuron Disorders				
higher in *Smn*$^{2B/-}$	P07356.2	annexin-2	1.54	0.03
higher in *Smn*$^{2B/-}$	Q9Z0E0.1	neurochondrin	1.34	0.02
higher in *Smn*$^{2B/-}$	Q9CZD3.1	glycine-tRNA ligase	1.46	0.006
higher in *Smn*$^{2B/-}$	Q80 × 71.1	transmembrane protein 106b	1.46	0.008
Mitochondrial Dysfunction				
higher in *Smn*$^{2B/-}$	Q9R088.2	thymidine kinase 2	1.47	0.03
higher in *Smn*$^{2B/-}$	Q9WTP7.3	adenylate kinase 3	1.37	0.04
Cholesterol Synthesis and Bioenergetics				
higher in *Smn*$^{2B/-}$	Q8BLN5.2	lanosterol synthase	1.52	0.05
higher in *Smn*$^{2B/-}$	Q8K0C4.1	lanosterol demethylase	1.52	0.04
higher in *Smn*$^{2B/-}$	Q88844.2	isocitrate dehydrogenase (cytoplasmic)	1.41	0.02
lower in *Smn*$^{2B/-}$	P0C027.1	DIPP3[a]	1.47	0.05
Neurotransmitter and Protein Clearance				
lower in *Smn*$^{2B/-}$	Q11011.2	puromycin sensitive aminopeptidase	2.39	0.002
higher in *Smn*$^{2B/-}$	Q99KK7.2	dipeptyl peptidase 3	1.41	0.02
higher in *Smn*$^{2B/-}$	Q8JZP2.2	synapsin-3	1.35	0.04
higher in *Smn*$^{2B/-}$	Q9D4H8.2/Q9JLV5.1	cullin-2/3	1.33/1.40	0.01/0.02

[a]DIPP: Diphosphoinositol polyphosphate phosphohydrolase 3-alpha.

To investigate additional connections between the observed protein imbalances and canonical cellular pathways, we compared *DAVID* enrichment analyses between our list of differentially expressed proteins and the total dataset **[Figure 4.7]**. Most notable were the enrichment of proteins that function in mitochondria dynamics, protein clearance, and lipid metabolism **[Figure 4.7]**. Following enrichment analysis, we were interested in delineating putative molecular regulators within these functional categories that may play important roles in the etiology of the SMA phenotype.

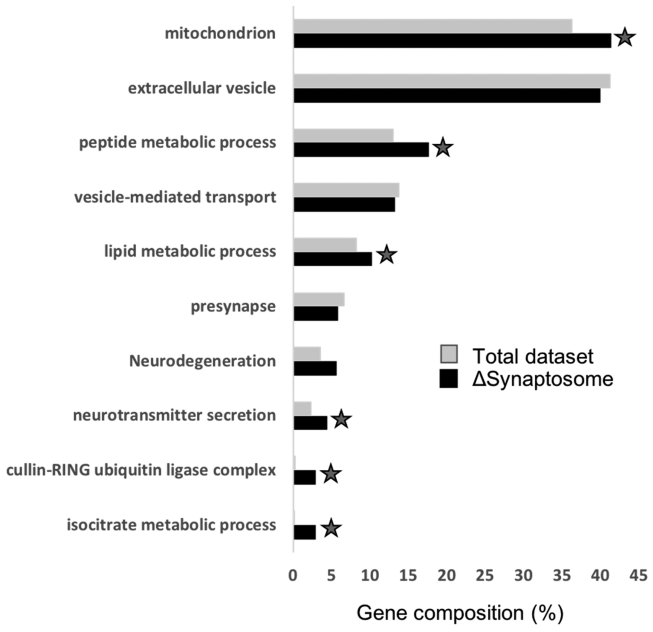


Figure 4.7 - Gene Ontology of differentially abundant synaptosome proteins. Qualitative comparison of DAVID enrichment analyses of synaptosomal proteins found to be differentially expressed (black bars, Δsynaptosome) and the total synaptosome proteome data set (gray bars, total data set). Gene Ontology mapped for proteins in relation to molecular function. Gene Ontology terms are shown on the y-axis with corresponding percent gene composition (x-axis). Notable differences are highlighted by stars.

Mitochondria Dysfunction: A commonality among many neuromuscular disorders is the disruption of mitochondrial homeostasis and bioenergetics [41]. However, the specific mitochondria pathways which are affected can differ substantially between these disorders. In SMA, the number of viable mitochondria within motor neurons has been shown to be severely diminished in response to decreased levels of SMN protein [32]. Our dataset found >85% of mitochondrial proteins to be at similar levels, suggesting no significant difference in the total number of intra-synaptic mitochondria between *Smn*$^{2B/-}$ and control mice. However, there were specific mitochondrial proteins whose levels were significantly different between *Smn*$^{2B/-}$ and control mice, suggesting that mitochondrial dysfunction may contribute to the degeneration of the synaptic terminals during the early onset of the disease. An interesting example was the observation of two enzymes active in mitochondrial nucleotide homeostasis, which were found to be higher in *Smn*$^{2B/-}$ mice; thymidine kinase 2 (TK2) (mtDNA synthesis) and adenylate kinase 3 (AK3) (ADP/ATP exchange) **[Table 4.2]** [42,43]. The TK2 result is surprising since it has been reported that TK2 deficiency, in contrast to the elevated levels seen in our dataset, leads to myopathy and mitochondrial toxicity [44]. Moreover, AK3 was specifically higher in *Smn*$^{2B/-}$ mice, while isozymes AK1 and AK2 were both found to be at equal levels between the two conditions **[Table S4.1]**. While AK3 is somewhat understudied, it is specifically localized to the mitochondrial matrix, which is the site of central metabolic processes such as fatty acid oxidation and the citric acid cycle [45-47]. In contrast, AK1 and AK2 are predominately localized to the cytosol and inner mitochondrial membrane respectively [48]. However, further studies will be required to determine the functional consequence of

overexpressing TK2 and AK3 with respect to synaptic mitochondria dysfunction and motor neuron survival.

Cholesterol Synthesis and Bioenergetics: Our current study has revealed intriguing alterations in proteins involved in the synthesis of metabolic second messengers and cholesterol biogenesis **[Table 4.2]**. Only 7 proteins were identified to be deficient in the $Smn^{2B/-}$ sample, one of them being DIPP-3α, a protein known to hydrolyze diphosphoinositol polyphosphates [49]. Although the biological activities of diphosphoinositol polyphosphates are poorly understood, their levels are tightly coordinated with the bioenergetic status of the cell [49]. Furthermore, diphosphoinositol polyphosphate accumulation has been shown to compete for plekstrin homology domain containing proteins and attenuate phosphatidylinositol (3,4,5) triphosphate (PIP3) signaling [50]. As PIP3 has important roles in postsynaptic processes and synaptic plasticity, it would be interesting to test if diphosphoinositol polyphosphate accumulation occurs in $Smn^{2B/-}$ mice due to lower DIPP3α levels, and if there is any subsequent effect on PIP3 signaling at the synapse of motor neurons [51].

We also observed a cluster of proteins at higher abundance ratios in the $Smn^{2B/-}$ samples that function in the production of cholesterol. Most notable was the detection of lanosterol synthase and lanosterol demethylase at significantly higher abundance levels in the $Smn^{2B/-}$ sample **[Table 4.2]**. Lanosterol synthase and lanosterol demethylase function in the biosynthesis of cholesterol and other sterols [52]. Moreover, a cytosolic isoform of isocitrate dehydrogenase (IDHc) was upregulated in $Smn^{2B/-}$ mice **[Table 4.2]**. IDHc is indirectly involved in the synthesis of cholesterol through the production of NADPH, a cofactor required by lanosterol demethylase [52,53]. Finally, cytoplasmic aconitase was

also detected at statistically higher levels in the *Smn*$^{2B/-}$ sample **[Table S4.1]**. Cytoplasmic aconitase catalyzes the interconversion of citrate to isocitrate and thus regulates the amount of NADPH generated from isocitrate by isocitrate dehydrogenase [54]. The role of cholesterol in motor neuron biology is quite dynamic. On the one hand, cholesterol has been well characterized as being a critical factor for synaptic formation and function [55]. On the other hand, cholesterol accumulation has also been implicated in oxidative stress induced cell death of motor neurons in ALS [56]. Collectively, the increase of four cholesterol regulating enzymes in the SMA sample encourages further investigation of the role of cholesterol in the pathology of SMA.

Neurotransmitter and Protein Clearance: Our proteomic investigation indicated alterations of a number of enzymes involved in protein turnover that may contribute to neurotoxicity dependent cell death in SMA pathology. For instance, we found that the protein abundance of dipeptidyl peptidase 3 (also known as enkephalinase B) and synapsin-3 to be increased in *Smn*$^{2B/-}$ mice **[Table 4.2]**. At the synapse, dipeptidyl peptidase 3 inactivates several critical neuropeptides and circulating hormones,[57] while synapsin-3 is involved in synaptic vesicle recycling and regulation of neurotransmitter release [57-59]. The identification of three members of the Cullin family of E3 ligases was intriguing due to the recent report of Cullin induced degradation of SMN protein [60]. Cullin-2 and Cullin-3 were both found to be elevated in the *Smn*$^{2B/-}$ sample while Cullin-5 levels were unaltered **[Table S4.1]**. Since motor neuron survival is critically dependent on SMN levels, Cullin-2 and Cullin-3 represent potential targets for elevating SMN levels and preventing motor neuron cell death [25]. Finally, within our dataset, the protein that showed the largest abundance ratio difference along with superior statistical significance

was the puromycin aminopeptidase (PSA) **[Table 4.2]**. PSA belongs to the M1 family of metallopeptidases and is active in neuronal tissues [61,62]. PSA has been characterized as a strong attenuator of proteotoxicity and is proposed to promote autophagy-mediated protein clearance of several known proteotoxins including polyQ-expanded huntingtin, ataxin, mutant alpha-synuclein, and tau [63,64]. With PSA being found to be over 2-fold lower in the synapse of motor neurons isolated from $Smn^{2B/-}$ mice, PSA may serve as a promising SMA synaptic biomarker. Also, mechanisms that elevate PSA protein levels should be investigated to determine if they can contribute to synaptic terminal equilibrium and motor neuron survival during the early onset of the disease.

4.4 - Conclusion

Although SMA is largely considered a motor neuron disease, recent evidence has shown defects in neural circuits that impact central synapses [10,12-15]. Indeed, the response of motor neurons to stimulation of afferent fibers is significantly decreased in SMA mice, even prior to onset of disease. However, the nature of molecular changes at the central synapses of SMA mice remains an open question. Here, our quantitative proteomic efforts have revealed novel protein alterations occurring within the central synapse of spinal cords from $Smn^{2B/-}$ mice at the pre-symptomatic stage. Intriguing changes involved imbalances of proteins that may contribute to neurotoxicity dependent cell death and synaptic dysfunction, encouraging new avenues of research into the role of synaptic homeostasis in SMA pathology.

4.5 - References

[1] J. Pearn, Incidence, prevalence, and gene frequency studies of chronic childhood spinal muscular atrophy, J Med Genet 15 (1978) 409-413.

[2] S. Lefebvre, L. Bürglen, S. Reboullet, O. Clermont, P. Burlet, L. Viollet, B. Benichou, C. Cruaud, P. Millasseau, M. Zeviani, Identification and characterization of a spinal muscular atrophy-determining gene, Cell 80 (1995) 155-165.

[3] J. Pearn, Classification of spinal muscular atrophies, Lancet 1 (1980) 919-921.

[4] K. Zerres, S. Rudnik-Schöneborn, Natural history in proximal spinal muscular atrophy. Clinical analysis of 445 patients and suggestions for a modification of existing classifications, Arch Neurol 52 (1995) 518-523.

[5] M.A. Farrar, S.B. Park, S. Vucic, K.A. Carey, B.J. Turner, T.H. Gillingwater, K.J. Swoboda, M.C. Kiernan, Emerging therapies and challenges in spinal muscular atrophy, Ann Neurol 81 (2017) 355-368.

[6] J.E. Sleeman, P. Ajuh, A.I. Lamondm, snRNP protein expression enhances the formation of Cajal bodies containing p80-coilin and SMN, J Cell Sci 114 (2001) 4407-4419.

[7] D.D. Coovert, T.T Le, P.E. McAndrew, J. Strasswimmer, T.O Crawford, J.R Mendell, S.E. Coulson, E.J. Androphy, T.W. Prior, A.H Burghes, The survival motor neuron protein in spinal muscular atrophy, Hum Mol Genet 6 (1997) 1205-1214.

[8] U.R. Monani, M. Sendtner, D.D. Coovert, D.W. Parsons, C. Andreassi, T.T Le, S. Jablonka, B. Schrank, W. Rossoll, T.W. Prior, G.E. Morris, A.H. Burghes, The human centromeric survival motor neuron gene (SMN2) rescues embryonic lethality in Smn(-/-) mice and results in a mouse with spinal muscular atrophy, Hum Mol Genet 9 (2000) 333-339.

[9] T.O. Gavrilina, V.L. McGovern, E. Workman, T.O. Crawford, R.G. Gogliotti, C.J. DiDonato, U.R. Monani, G.E. Morris, A.H. Burghes, Neuronal SMN expression corrects spinal muscular atrophy in severe SMA mice while muscle-specific SMN expression has no phenotypic effect, Hum Mol Genet 17 (2008) 1063-1075.

[10] L.M. Murray, L.H. Comley, D. Thomson, N. Parkinson, K. Talbot, T.H. Gillingwater, Selective vulnerability of motor neurons and dissociation of pre- and post-synaptic pathology at the neuromuscular junction in mouse models of spinal muscular atrophy, Hum Mol Genet 17 (2008) 949-962.

[11] H.B. Park, S.M. Lee, J.S. Lee, M.S. Park, K.I. Park, R. Namgung, C. Lee, Survival analysis of spinal muscular atrophy type I, Korean J Pediatr 53 (2010) 965-970.

[12] E. Dachs, M. Hereu, L. Piedrafita, A. Casanovas, J. Calderó, J.E. Esquerda, Defective neuromuscular junction organization and postnatal myogenesis in mice with severe spinal muscular atrophy, J Neuropathol Exp Neurol 70 (2011) 444-461.

[13] L. Kong, X. Wang, D.W. Choe, M. Polley, B.G. Burnett, M. Bosch-Marcé, J.W. Griffin, M.M. Rich, C.J. Sumner, Impaired synaptic vesicle release and immaturity of neuromuscular junctions in spinal muscular atrophy mice, J Neurosci 29 (2009) 842-851.

[14] L. Torres-Benito, M.F. Neher, R. Cano, R. Ruiz, L. Tabares, SMN requirement for synaptic vesicle, active zone and microtubule postnatal organization in motor nerve terminals, PloS ONE 6 (2011).

[15] R. Martínez-Hernández, S. Bernal, E. Also-Rallo, L. Alías, M.J. Barceló, M. Hereu, J.E. Esquerda, E.F. Tizzano, Synaptic defects in type I spinal muscular atrophy in human development, *J Pathol* 229 (2013) 49-61.

[16] K.K.Y. Ling, M.Y. Lin, B. Zingg, Z. Feng, C.P. Ko, Synaptic Defects in the Spinal and Neuromuscular Circuitry in a Mouse Model of Spinal Muscular Atrophy, PLoS ONE 5 (2010).

[17] G.Z. Mentis, D. Blivis, W. Liu, E. Drobac, M.E. Crowder, L. Kong, F.J. Alvarez, C.J. Sumner, M.J. O'Donovan, Early functional impairment of sensory-motor connectivity in a mouse model of spinal muscular atrophy, Neuron 69 (2011) 453-467.

[18] W.L. Imlach, E.S. Beck, B.J. Choi, F. Lotti, L. Pellizzoni et al., SMN is required for sensory-motor circuit function in Drosophila, Cell 151 (2012) 427–439.

[19] T.M. Wishart, C.A. Mutsaers, M. Riessland, M.M. Reimer, G. Hunter et al., Dysregulation of ubiquitin homeostasis and β-catenin signaling promote spinal muscular atrophy, J Clin Invest 124 (2014) 1821-1834.

[20] M. Bowerman, A. Beauvais, C.L. Anderson, R. Kothary, Rho-kinase inactivation prolongs survival of an intermediate SMA mouse model, Hum Mol Genet 19 (2010) 1468-1478.

[21] M. Eshraghi, E. McFall, S. Gibeault, R. Kothary, Effect of genetic background on the phenotype of the Smn2B/- mouse model of spinal muscular atrophy, Hum Mol Genet 25 (2016) 4494-4506.

[22] P.R. Dunkley, P.E. Jarvie, P.J. Robinson, A rapid Percoll gradient procedure for preparation of synaptosomes, Nat Protoc 3 (2008) 1718-1728.

[23] R. Gombar, T.E. Pitcher, J.A. Lewis, J. Auld, P.O. Vacratsis, Proteomic characterization of seminal plasma from alternative reproductive tactics of Chinook salmon (Oncorhynchus tswatchysha), J Proteomics 22 (2017) 1-9.

[24] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources, Nature Protoc 4 (2009) 44-57.

[25] D.W. Huang, B.T. Sherman, R.A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, Nucleic Acids Res 37 (2009) 1-13.

[26] G.M. Shepherd, K.M. Harris, Three-dimensional structure and composition of CA3-->CA1 axons in rat hippocampal slices: implications for presynaptic connectivity and compartmentalization, J Neurosci 18 (1998) 8300-8310.

[27] T. Misgeld, T.L. Schwarz, Mitostasis in Neurons: Maintaining Mitochondria in an Extended Cellular Architecture, Neuron 96 (2017) 651–666.

[28] A. Steevensz, R. Gombar, R. Vergilino, M.E. Cristescu, P.O. Vacratsis, Proteomic Profile of Daphnia pulex using Data-Independent Acquisition Mass Spectrometry and Ion Mobility Separation, Proteomics 18 (2018).

[29] J.C. Silva, M.V. Gorenstein, G.Z. Li, J.P. Vissers, S.J. Geromanos, Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition, Mol Cell Proteomics 5 (2006) 144-156.

[30] J.C. Nelson, A.K. Stavoe, D.A. Colón-Ramos, The actin cytoskeleton in presynaptic assembly, Cell Adh Mig 7 (2013) 379-387.

[31] I.Schuiki, G. Daum, Phosphatidylserine decarboxylases, key enzymes of lipid metabolism, IUBMB Life 61 (2009) 151-162.

[32] H. Chaytow, Y.T. Huang, T.H. Gillingwater, K.M.E. Faller, The role of survival motor neuron protein (SMN) in protein homeostasis, Cell Mol Life Sci 75 (2018) 3877-3894.

[33] A. Antonellis, R.E. Ellsworth, N. Sambuughin, I. Puls, A. Abel et al., Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V, Am J Hum Genet 72 (2003) 1293-1299.

[34] K.L. Seburn, L.A. Nangle, G.A. Cox, P. Schimmel, R.W. Burgess, An active dominant mutation of glycyl-tRNA synthetase causes neuropathy in a Charcot-Marie-Tooth 2D mouse model, Neuron 51 (2006) 715-726.

[35] H. Suzuki, M. Matsuoka, The Lysosomal Trafficking Transmembrane Protein 106B Is Linked to Cell Death, J Biol Chem 291 (2016) 21448-21460.

[36] S.C. Ling, M. Polymenidou, D.W. Cleveland, Converging mechanisms in ALS and FTD, disrupted RNA and protein homeostasis, Neuron 79 (2013) 416-438.

[37] V. Gerke, C.E. Creutz, S.E. Moss, Annexins, linking Ca2+ signalling to membrane Dynamics, Nat Rev Mol Cell Biol 6 (2005) 449-461.

[38] A. Gauthier-Kemper, M. Suárez Alonso, F. Sündermann, B. Niewidok, M. Fernandez et al., Annexins A2 and A6 interact with the extreme N terminus of tau and thereby contribute to tau's axonal localization, J Biol Chem 293 (2018) 8065-8076.

[39] D. Shafey, J.G. Boyer, K. Bhanot, R. Kothary, Identification of novel interacting protein partners of SMN using tandem affinity purification, J Proteome Res 4 (2010) 1659-1669.

[40] L.W. Thompson, K.D. Morrison, S.L. Shirran, E.J.N. Groen, T.H. Gillingwater et al., Neurochondrin interacts with the SMN protein suggesting a novel mechanism for spinal muscular atrophy pathology, J Cell Sci 131 (2018).

[41] C.D. Katsetos, S. Koutzaki, J.J. Melvin, Mitochondrial dysfunction in Neuromuscular Disorders, Semin Pediatr Neurol 20 (2013) 202-215.

[42] X. Zhou, N. Solaroli, M. Bjerke, J.B. Stewart, B. Rozell, M. Johansson, A. Karlsson, Progressive loss of mitochondrial DNA in thymidine kinase 2-deficient mice, Hum Mol Genet 17 (2008) 2329–2335.

[43] Y. Zhang, H. Launay, F. Liu, R. Lebrun, B. Gontero, Interaction between adenylate kinase 3 and glyceraldehyde-3-phosphate dehydrogenase from Chlamydomonas reinhardtii, FEBS J 285 (2018) 2495-2503.

[44] C. Garone, R.W. Taylor, A. Nascimento, J. Poulton, C. Fratter et al., Retrospective natural history of thymidine kinase 2 deficiency, J Med Genet 55 (2018) 515-521.

[45] T. Noma, K. Fujisawa, Y. Yamashiro, M. Shinohara, A. Nakazawa, T. Gondo, T. Ishihara, K. Yoshinobu, Structure and expression of human mitochondrial adenylate kinase targeted to the mitochondrial matrix, Biochem J 358 (2001) 225-232.

[46] S. Eaton, K. Bartlett, M. Pourfarzam, Mammalian mitochondrial beta-oxidation, Biochem J 320 (1996) 345-357.

[47] L.D. Osellame, T.S. Blacker, M.R. Duchen, Cellular and molecular mechanisms of mitochondrial function, Best Pract Res Clin Endocrinol Metab 26 (2012) 711-723.

[48] C. Köhler, A. Gahm, T. Noma, A. Nakazawa, S. Orrenius, B. Zhivotovsky, Release of adenylate kinase 2 from the mitochondrial intermembrane space during apoptosis, FEBS Lett 447 (1999) 10-12.

[49] S.B. Shears, N.A. Gokhale, H. Wang, A. Zaremba, Diphosphoinositol polyphosphates, what are the mechanisms?, Adv Enzyme Regul 51 (2010) 13-25.

[50] H.R. Luo, Y.E. Huang, J.C. Chen, A. Saiardi, M. Iijima, K. Ye, Y. Huang, E. Nagata, P. Devreotes, S.H. Snyder, Inositol pyrophosphates mediate chemotaxis in Dictyostelium via pleckstrin homology domain-PtdIns(3,4,5)P3 interactions, Cell 114 (2003) 559-572.

[51] K.L. Arendt, M. Royo, M. Fernández-Monreal, S. Knafo, C.N. Petrok, J.R. Martens, J.A. Esteban, PIP3 controls synaptic function by maintaining AMPA receptor clustering at the postsynaptic membrane, Nat Neurosci 13 (2009) 36-44.

[52] W.D. Nes, Biosynthesis of cholesterol and other sterols, Chem Rev 11 (2011) 6423-6451.

[53] H.J. Koh, S.M. Lee, B.G. Son, S.H. Lee, Z.Y. Ryoo, K.T. Chang, J.W. Park, D.C. Park, D. B.J. Song, R.L. Veech, H. Song, T.L. Huh, Cytosolic NADP+-dependent isocitrate dehydrogenase plays a key role in lipid metabolism, J Biol Chem 279 (2004) 39968-39974.

[54] S.J. Lloyd, H. Lauble, G.S. Prasad, C.D. Stout, The mechanism of aconitase, 1.8 A resolution crystal structure of the S642A citrate complex, Protein Sci 8 (1998) 2655-2662.

[55] F.W. Pfrieger, Role of cholesterol in synapse formation and function, Biochim Biophys Acta 1610 (2003) 271-280.

[56] F. Schmitt, G. Hussain, L. Dupuis, J.P. Loeffler, A. Henriques, A plural role for lipids in motor neuron diseases, energy, signaling and structure, Front Cell Neurosci 8 (2014).

[57] P. Kumar, V. Reithofer, M. Reisinger, S. Wallner, T. Pavkov-Keller, P. Macheroux, K. Gruber, Substrate complexes of human dipeptidyl peptidase III reveal the mechanism of enzyme inhibition, Sci Rep 6 (2016).

[58] J. Feng, P. Chi, T.A. Blanpied, Y. Xu, A.M. Magarinos, A. Ferreira, R.H. Takahashi, H.T. Kao, B.S. McEwen, T.A. Ryan, G.J. Augustine, P. Greengard, Regulation of neurotransmitter release by synapsin III, J Neurosci 22 (2002) 4372-4380.

[59] B. Porton, W.C. Wetsel, H.T. Kao, Synapsin III, role in neuronal plasticity and disease, Semin Cell Devel Biol. 22 (2011) 416-424.

[60] N. Rodriguez-Muela, N. K. Litterman, E.M. Norabuena, J.L. Mull, M.J. Galazo et al., Single-Cell Analysis of SMN Reveals Its Broader Role in Neuromuscular Disease, Cell reports 18 (2017) 1484-1498.

[61] D.B. Constam, A.R. Tobler, A. Rensing-Ehl, I. Kemler, L.B. Hersh, A. Fontana, Puromycin-sensitive aminopeptidase. Sequence analysis, expression, and functional characterization, J Biol Chem 270 (1995) 26931-26939.

[62] S. McLellan, S.H. Dyer, G. Rodriguez, L.B. Hersh, Studies on the tissue distribution of the puromycin-sensitive enkephalin-degrading aminopeptidases, J Neurochem 51 (1988) 1552-1559.

[63] F.M. Menzies, R. Hourez, S. Imarisio, M. Raspe, O. Sadiq, D. Chandraratna, C. O'Kane, K.L. Rock, E. Reits, A.L. Goldberg, Puromycin-sensitive aminopeptidase protects against aggregation-prone proteins via autophagy, Hum Mol Genet 19 (2010) 4573-4586.

[64] L.C. Kudo, L. Parfenova, G. Ren, N. Vi, M. Hui, Z. Ma, K. Lau, M. Gray, F. Bardag-Gorce, M. Wiedau-Pazos, K.S. Hui, S.L. Karsten, Puromycin-sensitive aminopeptidase(PSA/NPEPPS) impedes development of neuropathology in hPSA/TAU(P301L) double-transgenic mice, Hum Mol Genet 20 (2011) 1820-1833.

CHAPTER 5- GENERAL CONCLUSIONS AND FUTURE WORK

5.1 - Improving Proteome Coverage: Utilizing Data-Independent Acquisition Mass Spectrometry Employing Orthogonal Ion Mobility Separation to Profile The Proteomes of Biological Systems

The complexity of the proteome is astonishing. With the variation in number of proteoforms estimated to reach levels of five orders of magnitude, it is not surprising that this factor alone can overwhelm proteomic investigations from characterizing the proteome in its totality [1]. What's more, the challenges in proteomics is further compounded by the fact that the physico-chemical properties of individual proteins can differ drastically between other proteins. Therefore, reliance on any single acquisition method to analyze proteomes are likely going to be incomplete [2]. However, these cruxes of proteomics are somewhat mitigated by the ever-increasing technological advancements made to mass spectrometry–based analytical platforms, as well as combining various separation techniques to reduce protein complexity [2,3]. Therefore, it is from these challenges that warrants the continued development of methods for proteomic applications.

In collaboration with Dr. Melania Cristescu's lab (University of McGill), we have applied data-independent acquisition mass spectrometry (MS^E) with ion mobility ($HDMS^E$) to qualitatively study the proteome of *Daphnia pulex*. Based on our efforts, we were successful in identifying approximately 400 novel proteins that have never been reported using other acquisition methods. Although this was not the first proteomic investigation aimed at studying this model organism, it is the first time that $HDMS^E$ was employed to examine the proteome of *Daphnia pulex*. Most notably, from our data collected, we observed an approximate 50% increase in sequence coverage and at least a

2-fold greater number of peptides identified when using HDMS^E compared to a strategy just using MS^E. Therefore, we have determined that using orthogonal ion mobility separation in conjunction with data-dependent acquisition mass spectrometry will be advantageous towards label-free quantitative proteomic analysis.

Daphnia is a versatile model organism used for a variety of targeted proteomic studies [4] which range from testing changes to their proteome under hypoxic conditions [5,6], microgravity [7], temperature [8,9], to the presence of predators or parasites [4, 10-13]. However, perhaps the most important implication of using these species in proteomic investigations is to evaluate the ecotoxicological status of freshwater ecosystems [14-16]. The task of ecotoxicological studies is to assess the environmental damage induced by anthropomorphic perturbations [17]. It is well known that *Daphnia* are keystone species for aquatic habitats as well as a sentinel for environmental stress. The eco-responsiveness of the *Daphnia* genome is believed to contribute to tolerating fluctuating abiotic and biotic challenges in the environment, resulting in wide distribution of phenotypic plasticity observed in *Daphnia* species [18]. In addition, it has been documented that this crustacean arthropod, including *Daphnia pulex*, have colonized the freshwater ecosystem of the Great-Lakes [19]. Although we did not utilize quantitative measures in our proteomic analysis of the *Daphnia pulex* proteome, the methods we developed later on employing label-free quantitative analysis using HDMS^E can be applied to this system to monitor aquatic environmental health and contribute to the preservation of freshwater resources. Specially, by focusing on proteomic assays that screen for biomarkers from *Daphnia pulex* that are affected by heavy metal toxicity in aquatic habitats, or even exposure from microplastics. As a result, these proteomic assays could have immense potential for industrial uses to

assess effects of mining, and other development projects effect on surrounding water sheds and ecosystems.

5.2 – Label-Free Quantitative Proteomics is a Robust and Viable Approach in Revealing Functional Networks Related to Disease Models and Evolutionary Adaptative Strategies

While genomic and transcriptomic endeavors can be useful at approximating the protein landscape, the information they provide is limiting on the basis that single gene products can yield many different protein isoforms, and that the abundance of mRNA will not always correlate to protein expression and their relative abundances [20]. This important fact underlies the notion that information pertaining to proteins can only be measured at the protein level [21]. However, the dynamic range of protein abundances, depending on the sample, can vary between six to ten orders of magnitude and can change in a spatial-temporal manner [22]. Thus, a major challenge in quantitative proteomics is to accurately measure these changes based on the available technologies. In the works presented in this thesis, we report the utility in utilizing HDMS^E with label-free quantitative strategies to monitor changes in protein concentrations in two separate biological systems.

5.2.1-Measuring Changes in Seminal Fluid Proteins of Chinook Salmon using Label-Free Quantitative Proteomics

Alternate reproductive tactics (ARTs), presented in the population of chinook salmon males, provided a novel opportunity to use quantitative proteomic applications to test and measure differences in protein content of the seminal plasma of the two competing males (i.e Hooknose and Jacks). In collaboration with Dr. Trevor Pitcher's lab (University of Windsor), we examined their hypothesis in which the sexual competition that occurs between males of different mating behaviours, presented by the individual tactics of chinook salmon males, would exhibit dramatic disparities in protein content in the fluid of seminal plasma. Although the chinook salmon protein database is still currently under-characterized, I was successful in developing a trained database to emulate chinook salmon seminal plasma proteins. Using bioinformatic analysis, we exploited the advantage of using cross-species proteomics of closely related *Teleost* fish to create a protein database for chinook salmon.

This was the first reported case using label-free quantitative analysis, alongside with $HDMS^E$, to measure and compare changes to the seminal plasma proteome content from a biological system that displays ARTs. Most importantly, we were able to observe significant changes to several proteins which may serve the purpose of mediating sperm motility and overall sperm-quality. One of the primary variables that affects sperm competition is overall sperm velocity [23]. One of the primary motives for initiating our study was previous work from the Pitcher laboratory that showed that Hooknose sperm velocity was significantly decreased after the addition of seminal fluid from Jack males, while the reciprocal experiment did not show a decrease to the velocity of Jack sperm [24,25]. To reflect back on our proteomic data, our efforts have uncovered that a

phosphodiesterase (PDE) – a protein that metabolizes cyclic AMP (cAMP) – was significantly upregulated in Jack seminal plasma compared to Hooknose seminal plasma. Since cAMP levels have been shown to mediate flagellar beating of the sperm tail (which is the main mechanism of propulsion), it warrants future studies to focus on the effects of PDE to chinook salmon sperm motility [26-28]. However, to test this theory would first require the quantitative examination of chinook salmon sperm proteome. In addition, our analysis has shown the presence of exosomes in chinook salmon seminal fluid. Although it was not shown from out dataset, PDEs are known to reside within exosome. Therefore, the direction of future work should also test for the proteomes of seminal plasma exosomes.

5.2.2 Monitoring Changes in Protein Abundances in Synaptosomes Mimicking SMA Pathology

Like most neuromuscular myopathies, the attenuation of synaptic transmission results from the destabilization of intracellular cargo shuttling within axons [29]. SMA is no exception, it has been shown that SMN1 loss of function is directly involved in the deregulation of actin dynamics [30]. In addition, actin is a major component of pre-synaptic terminals and responsible for organization, mobilization, and trafficking of synaptic vesicles [31]. It is believed that the desensitisation of neuromuscular junctions is caused by disruption of proper synaptic regulation and facilitates in the denervation observed in SMA [32]. While there are disease-modifying therapies that have been developed to target SMA, its efficiency is relegated to only certain SMA phenotypes [33]. Thus, it is important to target synaptic terminals for proteomic investigation to identify non-canonical pathways related to SMA pathology.

Our collaboration with Dr. Rashmi Kothary (University of Ottawa) have revealed possible novel synaptic targets for SMA. In addition, this is the first reported case where label-free quantitation using HDMS^E was performed on synaptosomes from a SMA mouse model. Some of the significant changes to the proteome has shown an impact to cellular networks related to mitochondrial dynamics, cholesterol biogenesis, and protein clearance. Although our analysis did not yield any physiological impact on SMA research, one of our most intriguing findings was the discovery of a substantial alteration of puromycin-sensitive aminopeptidase (PSA) in the synapse of SMA mice. With a fold-change almost three times greater in abundance in the control sample to the SMA sample, this was the first time that PSA was identified in relation to SMA studies and it will be interesting to determine its function in the progression of SMA and its potential use as a diagnostic marker. Future work should employ transgenic knock-out experiments to test if down-regulation of PSA in wild type mice can induce the phenotypic neurodegeneration observed in SMA. In addition, the development of targeted mass spectrometry methods for PSA would further SMA research to determine if PSA is altered in other SMA model systems including clinical patients. This could also lead to investigating if a enzyme replacement approach to return PSA to normal levels would be a viable therapeutic option.

5.3 Concluding Statement

A major challenge, and objective in the life-sciences has been the characterizing of living systems at the molecular level. However, as technology enhances our capacity to shift the modes of analysis towards a systematic scale, so too does the ability to advance our understanding of biological systems. As a proof of concept, we have observed this trend - in three independent cases in the following thesis - where mass spectrometry was shown to be an indispensable tool for the purposes of proteomic investigations. The corollary of which has made proteomics an essential medium for the characterizing of the fundamental biological process of protein expression (qualitative and quantitative), of the dynamics between inter and intra molecular protein networks, of the protein-protein interactions that form molecular complexes, of the spectrum of diverse function exhibited by individual proteins, and of the identification of biological targets for drug design and personalized medicine. In conclusion, continued exploration of dynamic changes made to the proteome, multiprotein networks, and post-translational modifications by means of proteomic analysis will provide a deeper vision of cellular life and the mechanisms that regulate its homeostasis.

5.4 References

[1] S.D. Patterson, R.H. Aebersold, Proteomics: the first decade and beyond, Nat Genet 33 (2003) 311-323.

[2] H. Hongzhan, H.D. Shukla, W. Cathy, S. Satya, Challenges and solutions in proteomics, Curr Genomics 8 (2007) 21–28.

[3] R. Aebersold, M. Mann, Mass spectrometry-based proteomics, Nature 13 (2003) 198-207.

[4] T. Fröhlich, G.J. Arnold, R. Fritsch, LC-MS/MS-based proteome profiling in *Daphnia pulex* and *Daphnia longicephala*: the *Daphnia pulex* genome database as a key for high throughput proteomics in *Daphnia*, BMC Genomics 10 (2009).

[5] B. Zeis, T. Lamkemeyer, R.J. Paul RJ, Acclimatory responses of the Daphnia pulex proteome to environmental changes. I. Chronic exposure to hypoxia affects the oxygen transport system and carbohydrate metabolism, BMC Physiol 9 (2009).

[6] B. Zeis, D. Becker, P. Gerke, M. Koch, R.J. Paul, Hypoxia-inducible haemoglobins of Daphnia pulex and their role in the response to acute and chronic temperature Increase, Biochim Biophys Acta 1834 (2013) 1704-1710.

[7] B. Trotter, K.A. Otte, K. Schoppmann, The influence of simulated microgravity on the proteome of *Daphnia magna*, NPJ Microgravity 1 (2015).

[8] D. Becker, Y. Reydelet, J.A. Lopez, The transcriptomic and proteomic responses of Daphnia pulex to changes in temperature and food supply comprise environment-specific and clone-specific elements, BMC Genomics 19 (2018).

[9] S. Schwerin, B. Zeis, T. Lamkemeyer, Acclimatory responses of the Daphnia pulex proteome to environmental changes. II. Chronic exposure to different temperatures (10 and 20 degrees C) mainly affects protein metabolism, BMC Physiol 9 (2009).

[10] K.A. Otte, T. Fröhlich, G.J. Arnold, C. Laforsch, Proteomic analysis of Daphnia magna hints at molecular pathways involved in defensive plastic responses, BMC Genomics 15 (2014).

[11] K.A. Otte, I. Schrank, T. Fröhlich, G.J. Arnold, C. Laforsch, Interclonal proteomic responses to predator exposure in Daphnia magna may depend on predator composition of habitats, Mol Ecol 24 (2015) 3901-3917.

[12] Y. Lu, P.R. Johnston, S.R. Dennis, *Daphnia galeata* responds to the exposure to an ichthyosporean gut parasite by down-regulation of immunity and lipid metabolism, BMC Genomics 19 (2018).

[13] A.P. Cuco, N. Abrantes, F. Gonçalves, J. Wolinska, B.B. Castro BB, Interplay between fungicides and parasites: Tebuconazole, but not copper, suppresses infection in a *Daphnia-Metschnikowia* experimental model, PLOS ONE 12 (2017).

[14] H.J. Kim, P. Koedrith, Y.R. Seo, Ecotoxicogenomic approaches for understanding molecular mechanisms of environmental chemical toxicity using aquatic invertebrate, Daphnia model organism, Int J Mol Sci 16 (2015) 12261–12287

[15] M. Borgatta, C, Hernandez, L.A. Decosterd, N. Chèvre, P. Waridel, Shotgun ecotoxicoproteomics of Daphnia pulex: biochemical effects of the anticancer drug tamoxifen, J Proteome Res 2 (2015) 279-291.

[16] Bang, S.H., Hong, N., Ahn, J. *et al.* Proteomic analysis of *Daphnia magna* exposed to caffeine, ibuprofen, aspirin and tetracycline. Toxicol Environ. Health Sci. 7 (2015) 97–104.

[17] Relyea R, Hoverman J. Assessing the ecology in ecotoxicology: a review and synthesis in freshwater systems, Ecol Lett, 9 (2006) 1157-1171.

[18] J.K. Colbourne, M.E. Pfrender, D. Gilbert et al., The ecoresponsive genome of Daphnia pulex, Science 4 (2011) 555-561.

[19] T.S. Bridges, R.B. Wright, B.R. Gray, A.B. Gibson, T.M. Dillon, Chronic toxicity of Great Lakes sediments to Daphnia magna: elutriate effects on survival, reproduction and population growth, Ecotoxicology 5 (1996) 83-102.

[20] C. Manzoni, D.A. Kia, J. Vandrovcova, J. Hardy, N.W. Wood, P.A. Lewis, R. Ferrari, Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences, Brief Bioinform 1 (2018) 286-302.

[21] E.E. Gulcicek, C.M. Colangelo, W. McMurray et al., Proteomics and the analysis of proteomic data: an overview of current protein-profiling technologies, Curr Protoc Bioinformatics (2005).

[22] R.A. Zubarev, The challenge of the proteome dynamic range and its implications for in-depth proteomics, Proteomics 13 (2013) 723-726.

[23] M.J. Gage, C.P. Macfarlane, S. Yeates, R.G. Ward, J.B. Searle, G.A. Parker, Spermatozoal traits and sperm competition in Atlantic salmon: relative sperm velocity is the primary determinant of fertilization success, Curr Biol 6 (2004) 44-47.

[24] E.W. Flannery, I.A.E. Butts, M. Słowińska, A. Ciereszko, T.E. Pitcher, Alternative Reproductive Tactics in Chinook Salmon, Biol J Linn Soc Lond 108 (2013) 99-108.

[25] J.A. Lewis, T.E. Pitcher, The effects of rival seminal plasma on sperm velocity in the alternative reproductive tactics of Chinook salmon, Theriogenology 1 (2017) 24-29.

[26] L. Maréchal, C. Guillemette, S. Goupil, P. Blondin, P. Leclerc, F.J Richard, Cyclic nucleotide phosphodiesterases in human spermatozoa and seminal fluid: Presence of an active PDE10A in human spermatozoa, Biochim Biophys Acta Gen Subj 1861 (2017) 147-156.

[27] A. Bergeron, J.P. Aragon, C. Guillemette, A. Hébert, R. Sullivan, P. Blondin, F.J. Richard, Characterization of cAMP-phosphodiesterase activity in bovine seminal plasma, Andrology 4 (2016) 1123-1130.

[28] M. Balbach, V. Beckert, J.N. Hansen, D. Wachten, Shedding light on the role of cAMP in mammalian sperm physiology, Mol Cell Endocrinol 15 (2018) 111-120.

[29] W. Rossoll, G.J. Bassell, Spinal muscular atrophy and a model for survival of motor neuron protein function in axonal ribonucleoprotein complexes, Results Probl Cell Differ 28 (2009) 289-326.

[30] M. Bowerman, C.L. Anderson, A. Beauvais, P.P Boyl, W. Witke, R. Kothary, SMN, profilin IIa and plastin 3: a link between the deregulation of actin dynamics and SMA pathogenesis, Mol Cell Neurosci 42 (2009) 66-74.

[31] J.C. Nelson, A.K. Stavoe, D.A. Colón-Ramos, The actin cytoskeleton in presynaptic assembly, Cell Adh Migr 7 (2013) 379–387.

[32] K.K. Ling, R.M. Gibbs, Z. Feng, C.P. Ko, Severe neuromuscular denervation of clinically relevant muscles in a mouse model of spinal muscular atrophy, Hum Mol Genet 1 (2012) 185-195.

[33] E.J.N. Groen , K. Talbot, T.H. Gillingwater, Advances in therapy for spinal muscular atrophy: promises and challenges, Nat Rev Neurol 14 (2018) 214-224.

PERMISSIONS

Written permission for all copy-right information within this dissertation has been granted by the appropriate agencies. Material has been cited below in order as it appears in the dissertation followed by copies of the written permission.

Permission was granted by Elsevier to recreate the image for **Figure 1.1 and Figure 1.4** in Chapter 1 as follows:

Figure 1.1 - R.E. March, J.F.J. Todd, Radio frequency quadrupole technology: Evolution and contributions to mass spectrometry, International Journal of Mass Spectrometry 377 (2015) 316-328.

Figure 1.4 - J. Ponthus, E. Riches, Evaluating the multiple benefits offered by ion mobility-mass spectrometry in oil and petroleum analysis, Int. J. Ion Mobil. Spec 16 (2013) 95–103.

Research in Chapter 2 was originally published in Proteomics. ‡Steevensz A, ‡Gombar R, Vergilino R, Cristescu ME, Vacratsis PO., Proteomic Profile of Daphnia pulex using Data-Independent Acquisition Mass Spectrometry and Ion Mobility Separation. *Proteomics*. (**2018**)18(16). © 2018 Wiley & Sons.

Research in Chapter 3 was originally published in the Journal of Proteomics. Gombar R, Pitcher TE, Lewis JA, Auld J, Vacratsis PO., Proteomic characterization of seminal plasma from alternative reproductive tactics of Chinook salmon (Oncorhynchus tswatchysha). *J Proteomics*. 22;157 (**2017**) 1-9. © Elsevier

Research in chapter 4 was orginally published in the Journal of Proteome Research. ‡Eshraghi M, ‡Gombar R, De Repentigny Y, Vacratsis PO, Kothary R., Pathologic Alterations in the Proteome of Synaptosomes from a Mouse Model of Spinal Muscular Atrophy. *J Proteome Res*. 2;18(8) (**2019**) 3042-3051. © ACS Publications.

VITA AUCTORIS

NAME:	Robert Gombar
PLACE OF BIRTH:	Windsor, ON
YEAR OF BIRTH:	1991
EDUCATION:	Kennedy Collegiate High School, Windsor, ON, 2009
	University of Windsor, B.Sc. Honours Biological Sciences with thesis, Windsor, ON, 2013
	University of Windsor, PhD. Biochemistry, Windsor, ON, 2020